

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 November 2001 (08.11.2001)

PCT

(10) International Publication Number
WO 01/83729 A2

- (51) International Patent Classification⁷: C12N 15/00 (74) Agent: BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent & Trademark Dept., CH-4002 Basel (CH).
- (21) International Application Number: PCT/EP01/04863
- (22) International Filing Date: 30 April 2001 (30.04.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 09/562,934 1 May 2000 (01.05.2000) US
- (71) Applicants: NOVARTIS AG [CH/CH]; Schwarzwalddallee 215, CH-4058 Basel (CH). THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

- (71) Applicants and
(72) Inventors: NEMEROW, Glen, R. [US/US]; 462 Cerro Street, Encinitas, CA 92024 (US). VON SEGGERN, Daniel, J. [US/US]; Apartment 30, 5175 Luigi Terrace, San Diego, CA 92122 (US). FRIEDLANDER, Marty [US/US]; 1720 Zapo Street, Del Mar, CA 92014 (US).

WO 01/83729 A2

(54) Title: VECTORS FOR OCULAR TRANSDUCTION AND USE THEREOF FOR GENETIC THERAPY

(57) Abstract: Adenovirus vector-based gene therapy methods for treating ocular disorders are provided. Adenovirus vectors for therapy of ocular diseases and methods of treatment using the vectors are provided. Compositions, kits, and methods of preparation and use of the vectors for gene therapy are provided.

VECTORS FOR OCULAR TRANSDUCTION AND USE THEREOF FOR GENETIC THERAPY

Work described herein was supported by NIH grants EY11431 and HL54352. The government has certain rights in such subject matter.

5 RELATED APPLICATIONS

This application claims the benefit of priority to U.S. application Serial No. 09/562,934, filed May 1, 2000, to Glen R. Nemerow, Daniel Von Seggern,; Martin Friedlander, entitled "VECTORS FOR OCULAR TRANSDUCTION AND USE THEREFOR FOR GENETIC THERAPY".

- 10 This application is related to copending U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265, filed January 14, 2000)), to Daniel Von Seggern, Glen R. Nemerow, Paul Hallenbeck, Susan Stevenson, Yelena Skripchenko, filed January 14, 2000, entitled "Adenovirus Vectors, Packaging Cell Lines, Compositions, and
- 15 Methods for Preparation and Use," which is a continuation-in-part of U.S. Application 09/423,783 filed November 12, 1999 and claims the benefit of the filing date of U.S. Provisional Application 60/115,920 filed January 14, 1999. Where permitted, the contents and subject matter of each application and of the provisional application are incorporated in their entirety herein by reference.

20 FIELD OF INVENTION

- The present invention relates to gene therapy, especially to adenovirus vector-based gene therapy. In particular, adenovirus vectors for therapy of ocular diseases and methods of treatment using the vectors are provided. Compositions, kits, and methods of preparation and use of the vectors for gene
- 25 therapy are provided.

BACKGROUND OF THE INVENTION

Retinal dystrophies

- The eye is susceptible to a number of hereditary and/or age related degenerative disorders. In the United States, common causes of irreversible
- 30 blindness or severe loss of vision are retinal dystrophies (see, *e.g.*, Cotlier *et al.* (1995) *Surv. Ophthalmology* 40:51-61; Bird (1995) *Am. J. Ophthal.* 119: 543-562; and Adler (1996) *Arch Ophthal* 114:79-83). The retina is the sensory

-2-

tunic of the eye, containing light sensitive receptors, a complex of neurons, and pigmented epithelium, arranged in discrete layers. In humans, the macula is the portion of the retina that lies directly behind the lens. Cones, the photoreceptor cells responsible for central vision, are heavily concentrated in the macula.

5 Central dystrophies, which affect the macula, include Best's disease, age-related macular degeneration, and Stargardt's macular dystrophy. The peripheral retina is composed mainly of rods, which are responsible for side and night vision. Peripheral degenerative retinal diseases include retinitis pigmentosa, choroidemia and Bietti's crystalline dystrophy.

10 Macular degenerations are a heterogenous group of diseases, characterized by progressive central vision loss and degeneration of the macula and underlying retinal pigmented epithelium. Age-related macular degeneration (ARMD) is the most common form of the disease, affecting an estimated 20% of persons over 75 years of age. ARMD is poorly understood in terms of etiology
15 and pathogenesis. The very late onset of the disease has made genetic mapping particularly difficult. Certain macular degenerative conditions with a clear genetic basis, such as Stargardt's and Best's diseases, share many features with ARMD, but have been more amenable to molecular and genetic analysis.

Hereditary peripheral retinopathies are also relatively common. Retinitis
20 pigmentosa (RP), for example, affects approximately 1.5 million people worldwide. Substantial genetic heterogeneity has been observed in this condition, with over 20 chromosomal loci identified. A predisposition to retinitis pigmentosa can be inherited by autosomal dominant, autosomal recessive, X-linked or digenic mode. Mutations have been identified in seven genes, four of
25 which encode proteins in the rod phototransduction cascade: rhodopsin, alpha and beta subunits of rod cGMP phosphodiesterase, and rod cGMP cation-gated channel protein .alpha. subunit. Mutations in the peripherin/RDS gene have been linked to retinitis pigmentosa and macular degeneration. A single peripherin/RDS mutation apparently caused retinitis pigmentosa, pattern dystrophy and fundus
30 flavimaculatus, in different family members.

-3-

In spite of causal heterogeneity, there is significant clinical similarity among RP subtypes. Common signs and symptoms include early electroretinographic abnormalities, ophthalmoscopic findings, and protracted, contiguous expansion of the ring-like scotoma toward the macula, leading to progressively worsening tunnel vision. A recent hypothesis is that active photoreceptor cell death, which is characteristic of these genetically distinct disorders, is mediated by a common induction of apoptosis. It may be possible to treat these conditions by the administration of agents that block induction of apoptosis in photoreceptors, such as neurotrophic factors.

10 Adenovirus delivery vectors

Adenovirus, which is a DNA virus with a 36 kilobase (kb) genome, is very well-characterized and its genetics and genetic organization are understood. The genetic organization of adenoviruses permits substitution of large fragments of viral DNA with foreign DNA. In addition, recombinant adenoviruses are structurally stable and no rearranged viruses are observed after extensive amplification.

Adenoviruses have been employed as delivery vehicles for introducing desired genes into eukaryotic cells. The adenovirus delivers such genes to eukaryotic cells by binding to cellular receptors followed by internalization. The adenovirus fiber protein is responsible for binding to cells. The fiber protein has two domains, a rod-like shaft portion and a globular head portion that contains the receptor binding region. The fiber spike is a homotrimer, and there are 12 spikes per virion. Human adenoviruses bind to and infect a broad range of cultured cell lines and primary tissues from different species.

25 The 35,000+ base pair (bp) genome of adenovirus type 2 has been sequenced and the predicted amino acid sequences of the major coat proteins (hexon, fiber and penton base) have been described (see, *e.g.*, Neumann *et al.*, *Gene* 69: 153-157 (1988); Herisse *et al.*, *Nuc. Acids Res.* 9: 4023-4041 (1981); Roberts *et al.*, *J. Biol. Chem.* 259: 13968-13975 (1984); Kinloch *et al.*, *J. Biol. Chem.* 259: 6431-6436 (1984); and Chroboczek *et al.*, *Viol.* 161: 549-554, 1987).

-4-

The 35,935 bp sequence of Ad5 DNA is also known and portions of many other adenovirus genomes have been sequenced. The upper packaging limit for adenovirus virions is about 105% of the wild-type genome length (see, *e.g.*, Bett, *et al.*, *J. Virol.* 67(10): 5911-21, 1993). Thus, for Ad2 and Ad5, this
5 would be an upper packaging limit of about 38kb of DNA.

Adenovirus DNA also includes inverted terminal repeat sequences (ITRs) ranging in size from about 100 to 150 bp, depending on the serotype. The inverted repeats permit single strands of viral DNA to circularize by base-pairing of their terminal sequences to form base-paired "panhandle" structures that are
10 required for replication of the viral DNA.

For efficient packaging, the ITRs and the packaging signal (a few hundred bp in length) comprise the "minimum requirement" for replication and packaging of a genomic nucleic acid into an adenovirus particle. Helper-dependent vectors lacking all viral ORFs but including these essential *cis* elements (the ITRs and
15 contiguous packaging sequence) have been constructed.

Ad vectors have several distinct advantages as gene delivery vehicles. For example, recombination of such vectors is rare; there are no known associations of human malignancies with adenoviral infections despite common human infection with adenoviruses; the genome may be manipulated to
20 accommodate foreign genes of a fairly substantial size; and host proliferation is not required for expression of adenoviral proteins. Adenovirus (Ad)-based gene delivery vectors efficiently infect many different cells and tissues. This broad tropism, however, means that gene delivery cannot be directed to a specific target cell. A large fraction of intravenously administered adenovirus is
25 retained by the liver, which could lead to undesirable side-effects. Adenovirus may potentiate immune responses. For example, Adenovirus type 5 (Ad5) also transduces dendritic cells, which present antigens very efficiently, thereby possibly exacerbating the immune response against the vector. It has been proposed that vectors with different targeting efficiencies might eliminate these
30 problems, permitting a lower total particle dose and more specific targeting (see, *e.g.*, U.S. application Serial No. 09/482,682).

-5-

The wealth of information on adenovirus structure and mechanism of infection, its efficient infection of nondividing cells, and its large genetic capacity make adenovirus a popular gene therapy vector. The wide expression of receptors to which adenovirus binds makes targeting adenovirus vectors
5 difficult.

Hence there is a need to improve delivery and targeting of adenoviral vectors and also to provide treatments for ocular disorders. Therefore, it is an object herein to provide adenoviral vectors that specifically or selectively target cells in the eye. It is also an object herein to provide these vectors for treatment
10 of ocular disorders.

SUMMARY OF THE INVENTION

Degenerative ocular diseases, such as, but not limited to, retinitis pigmentosa, Stargardt's disease, diabetic retinopathies, retinal vascularization, and others (see, *e.g.*, Table below), have a genetic basis. Genes expressed in
15 the photoreceptor cells at the back of the retina are implicated in these diseases. Provided herein are recombinant viral vectors for targeting therapeutic products to these cells.

Recombinant adenoviral vectors that include nucleic acid that permits specific binding to these photoreceptors are provided. In particular, the vector
20 particles contain a fiber protein of Ad37 or a modified form thereof. As shown herein, fiber protein from Ad37 permits efficient infection of photoreceptor cells. Fiber proteins from other adenovirus D serotypes may also be used. In addition, the portions of the fiber protein, particularly those that interact with other viral structural proteins, such as penton, may be modified to resemble the viral source
25 of the other structural proteins. As exemplified herein, the recombinant virus provided herein include Ad5 structural components. The N-terminus of the Ad37 fiber protein, which interacts with the penton protein, is modified to resemble the Ad5 fiber protein N-terminus to ensure production of viral particles.

The recombinant adenoviral vectors are intended for gene therapy of
30 diseases in which genes expressed in the photoreceptors are implicated. Such diseases include, but are not limited to, degenerative ocular diseases, such as retinitis pigmentosa and Stargardt's disease. These vectors are also useful for

-6-

targeting to other ocular cells, such as conjunctival cells, which also bear receptors to which fiber from Ad37 and related serotypes bind.

The vectors will deliver therapeutic agents to the targeted cells for treatment of a variety of disorders (see *e.g.*, Tables 3 and 4, below)). The
5 therapeutic agents are intended for expression in the photoreceptors and for secretion from the photoreceptor cells, which are surrounded on one side by choroidal vasculature, and on the other side by retinal vasculature, thereby providing a means for delivery of products. In addition, expression of growth factors, such as VEGF and others, can be used to enhance blood flow to the
10 retina and prevent or slow the degeneration.

Therapeutic agents encoded by the recombinant adenoviral vectors include, but are not limited to, nucleic acid nucleic acid molecules encoding genes that are defective in certain hereditary disorders, nucleic acid molecules that encode antiangiogenics and antitumor agents for treatment of retinal
15 disorders, such as retinoblastomas; nucleic acid molecules encoding trophic factors, such as glial cell line-derived neurotrophic factor (GDNF) and ciliary neurotrophic factor (CNTF), growth factors and growth factor inhibitors, antiapoptotic factors, such as Bcl-2 (CNTF), antitumor agents, anti-angiogenics, and genes or portions thereof for gene replacement or repair of defective genes.
20 Hence, methods for treatment of inherited and acquired retinal diseases, including diseases involving neovascular and vascular degeneration are provided.

Methods for treating diseases involving genes expressed in photoreceptor cells are provided herein. The methods provided herein are practiced by
25 administration of the recombinant viral vectors by any means suitable for delivery to the photoreceptors. A preferred mode of administration is intraocular injection including intravitreal and subretinal injection. Other modes of administration include, but are not limited to, intrascleral, periorbital and intravenous administration. The vectors also can include photoreceptor-specific
30 promoters thereby providing a means, not only for specific targeting of expression in these cells, but also for photoreceptor-restricted transgene expression.

DETAILED DESCRIPTION OF THE INVENTION**A. DEFINITIONS**

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. All patents, applications, published applications and other publications and sequences from GenBank and other data bases referred to anywhere in the disclosure herein are incorporated by reference in their entirety.

As used herein, the amino acids, which occur in the various amino acid sequences appearing herein, are identified according to their three-letter or one-letter abbreviations. The nucleotides, which occur in the various DNA fragments, are designated with the standard single-letter designations used routinely in the art (see, Table 1).

As used herein, amino acid residue refers to an amino acid formed upon chemical digestion (hydrolysis) of a polypeptide at its peptide linkages. The amino acid residues described herein are preferably in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property is retained by the polypeptide. NH₂ refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxyl terminus of a polypeptide. In keeping with standard polypeptide nomenclature described in *J. Biol. Chem.*, 243:3552-59 (1969) and adopted at 37 C.F.R. § § 1.821 - 1.822, abbreviations for amino acid residues are shown in the following

Table:

Table 1
Table of Correspondence

SYMBOL		AMINO ACID
1-Letter	3-Letter	
Y	Tyr	tyrosine
G	Gly	glycine
F	Phe	phenylalanine
M	Met	methionine

-8-

SYMBOL		
A	Ala	alanine
S	Ser	serine
I	Ile	isoleucine
L	Leu	leucine
T	Thr	threonine
V	Val	valine
P	Pro	proline
K	Lys	lysine
H	His	histidine
Q	Gln	glutamine
E	Glu	glutamic acid
Z	Glx	Glu and/or Gln
W	Trp	tryptophan
R	Arg	arginine
D	Asp	aspartic acid
N	Asn	asparagine
B	Asx	Asn and/or Asp
C	Cys	cysteine
X	Xaa	Unknown or other

It should be noted that all amino acid residue sequences represented herein by formulae have a left to right orientation in the conventional direction of amino-terminus to carboxyl-terminus. In addition, the phrase "amino acid residue" is broadly defined to include the amino acids listed in the Table of Correspondence and modified and unusual amino acids, such as those referred to in 37 C.F.R. § § 1.821-1.822, and incorporated herein by reference. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or

-9-

more amino acid residues or to an amino-terminal group such as NH₂ or to a carboxyl-terminal group such as COOH.

In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and may be made generally without altering the biological activity of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, e.g., Watson et al., *Molecular Biology of the Gene*, 4th Edition, 1987, The Bejacmin/Cummings Pub. co., p.224).

Such substitutions are preferably made in accordance with those set forth in TABLE 2 as follows:

TABLE 2	
Original residue	Conservative substitution
Ala (A)	Gly; Ser
Arg (R)	Lys
Asn (N)	Gln; His
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
Gly (G)	Ala; Pro
His (H)	Asn; Gln
Ile (I)	Leu; Val
Leu (L)	Ile; Val
Lys (K)	Arg; Gln; Glu
Met (M)	Leu; Tyr; Ile
Phe (F)	Met; Leu; Tyr
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp; Phe
Val (V)	Ile; Leu

Other substitutions are also permissible and may be determined empirically or in accord with known conservative substitutions.

As used herein, a complementing plasmid describes plasmid vectors that deliver nucleic acids into a packaging cell line for stable integration into a chromosome in the cellular genome.

As used herein, a delivery plasmid is a plasmid vector that carries or delivers nucleic acids encoding a therapeutic gene or gene that encodes a

-10-

therapeutic product or a precursor thereof or a regulatory gene or other factor that results in a therapeutic effect when delivered *in vivo* in or into a cell line, such as, but not limited to a packaging cell line, to propagate therapeutic viral vectors.

5 As used herein, a variety of vectors with different requirements are described. For example, one vector is used to deliver particular nucleic acid molecules into a packaging cell line for stable integration into a chromosome. These types of vectors are generally identified herein as complementing
10 acid molecules in or into a cell line (*e.g.*, a packaging cell line) for the purpose of propagating therapeutic viral vectors; hence, these vectors are generally referred to herein as delivery plasmids. A third "type" of vector described herein is used to carry nucleic acid molecules encoding therapeutic proteins or polypeptides or regulatory proteins or are regulatory sequences to specific cells or cell types in a
15 subject in need of treatment; these vectors are generally identified herein as therapeutic viral vectors or recombinant adenoviral vectors or viral Ad-derived . vectors and are in the form of a virus particle encapsulating a viral nucleic acid containing an expression cassette for expressing the therapeutic gene.

 As used herein, a DNA or nucleic acid homolog refers to a nucleic acid
20 that includes a preselected conserved nucleotide sequence, such as a sequence encoding a therapeutic polypeptide. By the term "substantially homologous" is meant having at least 80%, preferably at least 90%, most preferably at least 95% homology therewith or a lesser percentage of homology or identity and conserved biological activity or function.

25 The terms "homology" and "identity" are often used interchangeably. In this regard, percent homology or identity may be determined, for example, by comparing sequence information using a GAP computer program. The GAP program utilizes the alignment method of Needleman and Wunsch (*J. Mol. Biol.* 48:443 (1970), as revised by Smith and Waterman (*Adv. Appl. Math.* 2:482
30 (1981). Briefly, the GAP program defines similarity as the number of aligned symbols (*i.e.*, nucleotides or amino acids) which are similar, divided by the total number of symbols in the shorter of the two sequences. The preferred default

-11-

parameters for the GAP program may include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) and the weighted comparison matrix of Gribskov and Burgess, *Nucl. Acids Res.* 14:6745 (1986), as described by Schwartz and Dayhoff, eds., *ATLAS OF PROTEIN SEQUENCE*
5 *AND STRUCTURE*, National Biomedical Research Foundation, pp. 353-358 (1979); (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

Whether any two nucleic acid molecules have nucleotide sequences that
10 are at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% "identical" can be determined using known computer algorithms such as the "FAST A" program, using for example, the default parameters as in Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 (1988). Alternatively the BLAST function of the National Center for Biotechnology Information database may be used to
15 determine identity.

In general, sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics*
20 *and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991).

25 While there exist a number of methods to measure identity between two polynucleotides or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo, H. & Lipton, D., *SIAM J Applied Math* 48:1073 (1988)). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge
30 Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H. & Lipton, D., *SIAM J Applied Math* 48:1073 (1988). Methods to determine identity and similarity are codified in computer programs. Preferred

-12-

computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J., *et al.*, *Nucleic Acids Research* 12(1):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F., *et al.*, *J Molec Biol* 215:403 (1990)).

5 Therefore, as used herein, the term "identity" represents a comparison between a test and a reference polypeptide or polynucleotide. For example, a test polypeptide may be defined as any polypeptide that is 90% or more identical to a reference polypeptide. As used herein, the term at least "90% identical to" refers to percent identities from 90 to 99.99 relative to the
10 reference polypeptides. Identity at a level of 90% or more is indicative of the fact that, assuming for exemplification purposes a test and reference polynucleotide length of 100 amino acids are compared. No more than 10% (i.e., 10 out of 100) amino acids in the test polypeptide differs from that of the reference polypeptides. Similar comparisons may be made between a test and
15 reference polynucleotides. Such differences may be represented as point mutations randomly distributed over the entire length of an amino acid sequence or they may be clustered in one or more locations of varying length up to the maximum allowable, e.g. 10/100 amino acid difference (approximately 90% identity). Differences are defined as nucleic acid or amino acid substitutions, or
20 deletions.

As used herein, genetic therapy involves the transfer of heterologous DNA to the certain cells, target cells, of a mammal, particularly a human, with a disorder or conditions for which such therapy is sought. The DNA is introduced into the selected target cells in a manner such that the heterologous DNA is
25 expressed and a therapeutic product encoded thereby is produced. Alternatively, the heterologous DNA may in some manner mediate expression of DNA that encodes the therapeutic product, it may encode a product, such as a peptide or RNA that in some manner mediates, directly or indirectly, expression of a therapeutic product. Genetic therapy may also be used to deliver nucleic
30 acid encoding a gene product to replace a defective gene or supplement a gene product produced by the mammal or the cell in which it is introduced. The introduced nucleic acid may encode a therapeutic compound, such as a growth

-13-

factor inhibitor thereof, or a tumor necrosis factor or inhibitor thereof, such as a receptor therefor, that is not normally produced in the mammalian host or that is not produced in therapeutically effective amounts or at a therapeutically useful time. The heterologous DNA encoding the therapeutic product may be modified
5 prior to introduction into the cells of the afflicted host in order to enhance or otherwise alter the product or expression thereof.

As used herein, heterologous DNA is DNA that encodes RNA and proteins that are not normally produced *in vivo* by the cell in which it is expressed or that mediates or encodes mediators that alter expression of endogenous DNA by
10 affecting transcription, translation, or other regulatable biochemical processes. Heterologous DNA may also be referred to as foreign DNA. Any DNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which it is expressed is herein encompassed by heterologous DNA. Examples of heterologous DNA include, but are not limited to, DNA that encodes
15 traceable marker proteins, such as a protein that confers drug resistance, DNA that encodes therapeutically effective substances, such as anti-cancer agents, enzymes and hormones, and DNA that encodes other types of proteins, such as antibodies. Antibodies that are encoded by heterologous DNA may be secreted or expressed on the surface of the cell in which the heterologous DNA has been
20 introduced.

Hence, herein heterologous DNA or foreign DNA, refers to a DNA molecule not present in the exact orientation and position as the counterpart DNA molecule found in the corresponding wild-type adenovirus. It may also refer to a DNA molecule from another organism or species (*i.e.*, exogenous) or
25 from another Ad serotype.

As used herein, a therapeutically effective product is a product that is encoded by heterologous DNA that, upon introduction of the DNA into a host, a product is expressed that effectively ameliorates or eliminates the symptoms, manifestations of an inherited or acquired disease or that cures said disease.

30 Typically, DNA encoding the desired heterologous DNA is cloned into a plasmid vector and introduced by routine methods, such as calcium-phosphate mediated DNA uptake (see, (1981) Somat. Cell. Mol. Genet. 7:603-616) or

-14-

microinjection, into producer cells, such as packaging cells. After amplification in producer cells, the vectors that contain the heterologous DNA are introduced into selected target cells.

As used herein, an expression or delivery vector refers to any plasmid or virus into which a foreign or heterologous DNA may be inserted for expression in a suitable host cell — *i.e.*, the protein or polypeptide encoded by the DNA is synthesized in the host cell's system. Vectors capable of directing the expression of DNA segments (genes) encoding one or more proteins are referred to herein as "expression vectors." Also included are vectors that allow cloning of cDNA (complementary DNA) from mRNAs produced using reverse transcriptase.

As used herein, a gene is a nucleic acid molecule whose nucleotide sequence encodes RNA or polypeptide. A gene can be either RNA or DNA. Genes may include regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

As used herein, tropism with reference to an adenovirus refers to the selective infectivity or binding that is conferred on the particle by the fiber protein, such as by the C-terminus portion that comprises the knob.

As used herein, isolated with reference to a nucleic acid molecule or polypeptide or other biomolecule means that the nucleic acid or polypeptide has separated from the genetic environment from which the polypeptide or nucleic acid were obtained. It may also mean altered from the natural state. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Thus, a polypeptide or polynucleotide produced and/or contained within a recombinant host cell is considered isolated. Also intended as an "isolated polypeptide" or an "isolated polynucleotide" are polypeptides or polynucleotides that have been purified, partially or substantially, from a recombinant host cell or from a native source. For example, a recombinantly produced version of a compound can be substantially purified by the one-step method described in

-15-

Smith and Johnson, *Gene* 67:31-40 (1988). The terms isolated and purified are sometimes used interchangeably.

Thus, by "isolated" is meant that the nucleic acid is free of the coding sequences of those genes that, in the naturally-occurring genome of the organism (if any) immediately flank the gene encoding the nucleic acid of interest. Isolated DNA may be single-stranded or double-stranded, and may be genomic DNA, cDNA, recombinant hybrid DNA, or synthetic DNA. It may be identical to a native DNA sequence, or may differ from such sequence by the deletion, addition, or substitution of one or more nucleotides.

Isolated or purified as it refers to preparations made from biological cells or hosts means any cell extract containing the indicated DNA or protein including a crude extract of the DNA or protein of interest. For example, in the case of a protein, a purified preparation can be obtained following an individual technique or a series of preparative or biochemical techniques and the DNA or protein of interest can be present at various degrees of purity in these preparations. The procedures may include for example, but are not limited to, ammonium sulfate fractionation, gel filtration, ion exchange chromatography, affinity chromatography, density gradient centrifugation and electrophoresis.

A preparation of DNA or protein that is "substantially pure" or "isolated" should be understood to mean a preparation free from naturally occurring materials with which such DNA or protein is normally associated in nature. "Essentially pure" should be understood to mean a "highly" purified preparation that contains at least 95% of the DNA or protein of interest.

A cell extract that contains the DNA or protein of interest should be understood to mean a homogenate preparation or cell-free preparation obtained from cells that express the protein or contain the DNA of interest. The term "cell extract" is intended to include culture media, especially spent culture media from which the cells have been removed.

As used herein, a packaging cell line is a cell line that provides a missing gene product or its equivalent.

As used herein, an adenovirus viral particle is the minimal structural or functional unit of a virus. A virus can refer to a single particle, a stock of

-16-

particles or a viral genome. The adenovirus (Ad) particle is relatively complex and may be resolved into various substructures.

As used herein, "penton" or "penton complex" are preferentially used herein to designate a complex of penton base and fiber. The term "penton" may also be used to indicate penton base, as well as penton complex. The meaning of the term "penton" alone should be clear from the context within which it is used.

As used herein, a plasmid refers to an autonomous self-replicating extrachromosomal circular nucleic acid molecule, typically DNA.

10 As used herein, a post-transcription regulatory element (PRE) is a regulatory element found in viral or cellular messenger RNA that is not spliced, i.e. intronless messages. Examples include, but are not limited to, human hepatitis virus, woodchuck hepatitis virus, the TK gene and mouse histone gene. The PRE may be placed before a polyA sequence and after a heterologous DNA
15 sequence.

As used herein, pseudotyping describes the production of adenoviral vectors having modified capsid protein or capsid proteins from a different serotype than the serotype of the vector itself. One example, is the production of an adenovirus 5 vector particle containing an Ad37 fiber protein. This may be
20 accomplished by producing the adenoviral vector in packaging cell lines expressing different fiber proteins.

As used herein, promoters of interest herein may be inducible or constitutive. Inducible promoters will initiate transcription only in the presence of an additional molecule; constitutive promoters do not require the presence of
25 any additional molecule to regulate gene expression. a regulatable or inducible promoter may also be described as a promoter where the rate or extent of RNA polymerase binding and initiation is modulated by external stimuli. Such stimuli include, but are not limited to various compounds or compositions, light, heat, stress and chemical energy sources. Inducible, suppressible and repressible
30 promoters are considered regulatable promoters. Preferred promoters herein, are promoters that are selectively expressed in ocular cells, particularly photoreceptor cells.

-17-

As used herein, receptor refers to a biologically active molecule that specifically or selectively binds to (or with) other molecules. The term "receptor protein" may be used to more specifically indicate the proteinaceous nature of a specific receptor.

5 As used herein, recombinant refers to any progeny formed as the result of genetic engineering. This may also be used to describe a virus formed by recombination of plasmids in a packaging cell.

As used herein, a transgene or therapeutic nucleic acid molecule includes DNA and RNA molecules encoding an RNA or polypeptide. Such molecules may
10 be "native" or naturally-derived sequences; they may also be "non-native" or "foreign" that are naturally- or recombinantly-derived. The term "transgene," which may be used interchangeably herein with the term "therapeutic nucleic acid molecule," is often used to describe a heterologous or foreign (exogenous) gene that is carried by a viral vector and transduced into a host cell.

15 Therefore, therapeutic nucleotide nucleic acid molecules include antisense sequences or nucleotide sequences which may be transcribed into antisense sequences. Therapeutic nucleotide sequences (or transgenes) all include nucleic acid molecules that function to produce a desired effect in the cell or cell nucleus into which said therapeutic sequences are delivered. For example, a therapeutic
20 nucleic acid molecule can include a sequence of nucleotides that encodes a functional protein intended for delivery into a cell which is unable to produce that functional protein.

As used herein, the vitreous of the eye refers to material that fills the chamber behind the lens of the eye (i.e., vitreous humor or vitreous body).

25 As used herein, a promoter region refers to the portion of DNA of a gene that controls transcription of the DNA to which it is operatively linked. The promoter region includes specific sequences of DNA that are sufficient for RNA polymerase recognition, binding and transcription initiation. This portion of the promoter region is referred to as the promoter. In addition, the promoter region
30 includes sequences that modulate this recognition, binding and transcription initiation activity of the RNA polymerase. These sequences may be *cis* acting or

-18-

may be responsive to *trans* acting factors. Promoters, depending upon the nature of the regulation, may be constitutive or regulated.

Thus, promoters are nucleic acid fragments that contain a DNA sequence that controls the expression of a gene located 3' or downstream of the promoter. The promoter is the DNA sequence to which RNA polymerase specifically binds and initiates RNA synthesis (transcription) of that gene, typically located 3' of the promoter. A promoter also includes DNA sequences that direct the initiation of transcription, including those to which RNA polymerase specifically binds. If more than one nucleic acid sequence encoding a particular polypeptide or protein is included in a therapeutic viral vector or nucleotide sequence, more than one promoter or enhancer element may be included, particularly if that would enhance efficiency of expression.

A regulatable or inducible promoter may be described as a promoter wherein the rate of RNA polymerase binding and initiation is modulated by external stimuli. (see, *e.g.*, U.S. Patent Nos. 5,750,396 and 5,998,205). Such stimuli include various compounds or compositions, light, heat, stress, chemical energy sources, and the like. Inducible, suppressible and repressible promoters are considered regulatable promoters.

Regulatable promoters may also include tissue-specific promoters. Tissue-specific promoters direct the expression of the gene to which they are operably linked to a specific cell type. Tissue-specific promoters cause the gene located 3' of it to be expressed predominantly, if not exclusively, in the specific cells where the promoter expressed its endogenous gene. Typically, it appears that if a tissue-specific promoter expresses the gene located 3' of it at all, then it is expressed appropriately in the correct cell types (see, *e.g.*, Palmiter et al. (1986) Ann. Rev. Genet. 20: 465-499).

As used herein, the phrase "operatively linked" generally means the sequences or segments have been covalently joined into one piece of DNA, whether in single or double stranded form, whereby control sequences on one segment control expression or replication or other such control of other segments. The two segments are not necessarily contiguous.

-19-

As used herein, exogenous encompasses any therapeutic composition that is administered by the therapeutic methods provided herein. Thus, exogenous may also be referred to herein as foreign, or non-native or other equivalent expression.

5 **B. Ad37 fiber tropism**

The adenovirus fiber protein is a major determinant of adenovirus tropism (Gall *et al.* (1996) *J. Virol.* 70:2116-2123; Stevenson *et al.* (1995) *J. Virol.* 69:2850-2857). The fiber protein extends from the capsid and mediates viral binding to the cell surface by binding to specific cell receptors (Philipson *et al.* 10 (1968) *J. Virol.* 2:1064-1075). The fiber is a trimeric protein that includes an N-terminal tail domain that interacts with the adenovirus penton base, a central shaft domain of varying length, and a C-terminal knob domain that contains the cell receptor binding site (Chroboczek *et al.* (1995) *Curr.Top.Microbiol.Immunol.* 15 199:163-200; Riurok *et al.* (1990) *J.Mol.Biol.* 215:589-596; Stevenson *et al.* (1995) *J. Virol.* 69:2850-2857). Fiber proteins of most adenovirus subgroups have been shown to bind specifically or selectively to the 46 kDa coxsackievirus-adenovirus receptor (CAR), (Bergelson *et al.* (1997) *Science* 275:1320-1323; Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915). CAR appears to be expressed in a variety of human tissues, including the lung, at various levels (Bergelson *et al.* 20 (1997) *Science* 275:1320-1323), but Ad37 binds poorly to lung epithelial cells (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). This suggests that the tropism of this serotype may be influenced by factors independent of CAR expression.

Structural and biochemical data also suggest that distinct receptor binding sites are located on different regions of the Ad5 and Ad37 fiber knobs.

25 Adopting the nomenclature of Xia *et al.* (Xia *et al.* (1994) *Structure* 2:1259-1270), the receptor binding site for Ad5 is located at the AB-loop on the side of the fiber knob (Bewley *et al.* (1999) *Science* 286:1579-1583; Roelvink *et al.* (1999) *Science* 286:1568-1571). It is known that a lysine residue at position 240 of the Ad37 fiber, located in the CD-loop, is important for receptor binding 30 (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). The co-crystal structure of the Ad12 knob and the N-terminal domain of CAR (Bewley *et al.* (1999) *Science* 286:1579-1583) show that the CD-loop does not contact CAR. It thus appears

-20-

that different regions of the Ad5 and Ad37 fiber knobs recognize distinct cell receptors.

A 46 kDa receptor for coxsackieviruses and adenoviruses (CAR) mediates attachment for many adenovirus serotypes. The wide distribution of CAR fails to explain why certain adenovirus serotypes (i.e. Ad37) are highly associated with severe ocular infections such as epidemic keratoconjunctivitis (EKC). Ad37 does not use CAR, but instead uses a glycoprotein that contains sialic acid as its primary receptor (Arnberg *et al.* ((2000) *J. Virol.* 74:42-48). The modest number of Ad37 binding sites per cell (Huang *et al.* (1999) *J. Virol.* 73:2798-2802) also suggests that Ad37 recognizes a specific glycoprotein as its primary receptor for binding to conjunctival cells.

Adenovirus type 37 (subgroup D) has been associated with infections of the eye and genital tract. The tropism of Ad37 derives from the binding preference of its fiber protein, which binds to a receptor located on the surface of cells including Chang C, conjunctival epithelial cell line (Huang *et al.* (1999) *J. Virology* 73:2798-2802).

A protein receptor that is preferentially expressed on conjunctival cells to which Ad37 fiber binds is shown herein. The preferential expression of the Ad37 receptor protein on conjunctival cells suggests that this receptor likely influences Ad37 tropism and should play a key role in ocular pathogenesis. It is shown herein that Ad37 uses a distinct protein receptor that is selectively expressed on conjunctival cells. It is shown that Ad37 binds well to conjunctival cells (Chang C), but poorly to lung carcinoma cells (A549). To determine if infection correlated with cell binding, an Ad5 vector containing the Ad37 fiber protein was constructed. The 'pseudotyped' vector delivered transgenes to Chang C cells better than to A549 cells. Ad37 binding was abolished by protease treatment of Chang C cells, indicating the receptor is a membrane protein. Ad37 binding to conjunctival cells is shown herein to be calcium-dependent. It is also shown that Ad37 infection was not inhibited by a function-blocking anti-CAR monoclonal antibody, which is a feature distinct from Ad5 fiber interaction with CAR. Using a virus overlay protein blot assay (VOPBA), calcium-dependent Ad37 binding to a 50 kDa membrane protein on Chang C

-21-

cells, but not A549 cells was detected. Ad19p a closely related serotype that fails to bind to conjunctival cells, does not recognize the 50 kDa protein. Together, these data indicate that the 50 kDa protein is a candidate receptor for Ad37 on conjunctival cells.

5 Significantly, it is also shown herein that, upon administration of the vector to the vitreous humor, the recombinant adenovirus with the Ad37 fiber preferentially and selectively binds to photoreceptor cells. Hence, a recombinant adenoviral delivery vehicle that has an Ad37 fiber protein can serve as a vector for delivery of therapeutic agents to the eye for treatment of ocular disorders,
10 including genetic and acquired disorders. The identification of the receptor for Ad37 and the resulting recognition of Ad37 tropism allows targeting of adenovirus vectors to specific human ocular cells.

As noted, fiber plays a crucial role in adenovirus infection by attaching the virus to a specific receptor on a cell surface. Hexon, penton and fiber
15 capsomeres are the major components on the surface of the virion. The fiber is an elongated protein which exists as a trimer of three identical polypeptides (polypeptide IV) of 582 amino acids in length. An adenovirus fiber includes three domains: an N-terminal tail domain that interacts with penton base; a shaft composed of variable numbers of repeats of a 15-amino-acid segment that
20 forms beta-sheet and beta-bends; and a knob at the C-terminus ("head domain") that contains the type-specific antigen and is responsible for binding to the cell surface receptor. The gene encoding the fiber protein from Ad2 has been expressed in human cells and has been shown to be correctly assembled into trimers, glycosylated and transported to the nucleus (see, *e.g.*, Hong and Engler,
25 *Virology* 185: 758-761, 1991). Thus, alteration of the fiber in recombinant Ad vectors can lead to alteration in gene delivery.

As shown herein, alteration of fiber in recombinant Ad vectors such that the fiber is derived from Ad37 or another adenovirus serotype D, provides a means for selective delivery of a recombinant virus to particular cells in the eye,
30 including conjunctival cells, and most significantly photoreceptors, thereby providing a means for targeted delivery to photoreceptor cells.

-22-

Photoreceptor cells are implicated in a number of hereditary and acquired retinal degenerative disorders. In addition, photoreceptor cells are located such that products produced therein can be delivered to other areas of the eye by virtue of the blood flow in the vicinity of the photoreceptor cells and also by virtue of the proximity of the photoreceptors to the retinal pigmented epithelium (RPE) and other retinal cells.

Hence it is contemplated herein that the recombinant viral vector will include a packaged recombinant adenovirus genome containing at least the minimal elements for replication and packaging; heterologous DNA encoding a desired gene product, typically a therapeutic product or plurality of products, such as several trophic factors, whose combined activity is effective for treating a disorder, such as a retinal degenerative disorder; and the resulting virion particles will include a fiber that has a sufficient portion to confer specific targeting to photoreceptor cells when the recombinant viral particles are introduced into the aqueous humor of a mammalian, preferably a human, eye, or otherwise contacted with the photoreceptor cells. The fiber may be a chimeric protein that has been modified for effective interaction with other coat structural proteins, such as penton. In addition, the fiber may be modified to include other elements that alter its tropism to permit binding to other cells as well (see, *e.g.*, U.S. Patent Nos. 5,756,086 and 5,543,328, International PCT application No. WO 95/26412 and WO 98/44121 and Krasnykh, *et al.* (*J. Virol.* 70: 6839-46, 1996).

C. Construction of the viral particles

1. Selection of viral genome and fiber protein

Methods for preparing recombinant adenoviral vectors for gene product delivery are well known. Preferred among those are the methods exemplified herein (see EXAMPLES) and also described in copending U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265, filed January 14, 2000, which claims priority to U.S. provisional application Serial No. 60/115,920, as does U.S. application Serial No. 09/482,682)).

-23-

As noted, any desired recombinant adenovirus is contemplated for use in the methods herein as long as the viral genome is packaged in a capsid that includes at least the portion of a fiber protein that provides selective binding to photoreceptor cells. This fiber protein is preferably from an adenovirus type D serotype and is preferably an Ad37 fiber. The fiber protein should retain the knob region at the C-terminus ("head domain") from the Ad virus of subgroup D that contains the type-specific antigen and is responsible for binding to the cell surface receptor. Hence the fiber protein can be a chimeric fiber protein as long as it retains a sufficient portion of the type D serotype to specifically or selectively bind to photoreceptor cells. Generally the portion retained will be all or a portion of the knob region. The precise amount of knob region required can be determined empirically by including portions thereof and identifying the minimum residues from and Ad type D serotype, preferably Ad37, to effect selective targeting of a virion packaged with such fiber to photoreceptors in the eye upon introduction of the packaged virion into the aqueous humor.

Recombinant adenovirus containing heterologous nucleic acids that encode a desired product, such a gene to correct a genetic defect, may be made by any methods known to those of skill in the art. The viruses must be packaged in a cell line that results in expression of fiber on the particles that specifically, electively or preferentially targets (binds and results in internalization) the viral particle to cells in the eye. The fiber protein from Ad37 and other Adenoviruses of serotype D that infect the eye effects such targeting. The resulting adenovirus particles that express such fiber is administered by intraocular injection, subretinal injection, particularly intravitreal injection, or any means that results in preferential accumulation in photoreceptor cells.

The family of Adenoviridae includes many members with at least 47 known serotypes of human adenovirus (Ad1-Ad47) (Shenk, *Virology*, Chapter 67, in Fields *et al.*, eds. Lippincott-Raven, Philadelphia, 1996,) as well as members of the genus Mastadenovirus including human, simian, bovine, equine, porcine, ovine, canine and opossum viruses and members of the Aviadenovirus genus, including bird viruses, such as CELO.

-24-

Thus it is contemplated that the methods herein can be applied to any recombinant viral vectors derived from any adenovirus species. One of skill in the art would have knowledge of the different adenoviruses (see, *e.g.*, Shenk, *Virology*, Chapter 67, in Fields *et al.*, eds. Lippincott-Raven, Philadelphia, 1996,) and can construct recombinant viruses containing portions of the genome of any such virus.

In the exemplified embodiment, viral particles with Ad37 fiber were prepared. Site-directed mutations were made to the Ad37 fiber gene to make the tail sequence more closely match that of Ad5 to facilitate Ad37 fiber binding to the Ad5 penton base. The plasmid for the expression of the Ad37 fiber protein, pDV80, contains the CMV promoter, the adenovirus type 5 tripartite leader (TPL), and the modified Ad37 fiber gene sequence. Genes of interest, such as nucleic acid encoding the β subunit of cGMP phosphodiesterase (β PDE), β -glucuronidase, rhodopsin, growth factors, anti-cancer agents, growth factor receptors and other anti-angiogenic agents, and anti-apoptotic agents, can be incorporated into these vectors using the methods known to those of skill in the art and exemplified herein.

Known adenovirus vectors, previously constructed for intraocular therapy (see, *e.g.*, Bennett *et al.* (1996) *Nature Medicine* 2:649-654, which provides an Ad virus encoding β PDE for treatment of retinitis pigmentosa; Cayouette *et al.* (1998) *Human Gene Therapy* 8:423-430, which provides an Ad vector that expresses CNTF for treatment of retinitis pigmentosa and other retinal degenerative diseases; and Li *et al.* (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92:7700-7704, which provides an Ad virus vector that encodes a human β -glucuronidase for treatment of lysosomal storage disease caused by β -glucuronidase deficiency) can be modified by repackaging the recombinant genome using a packaging line that expresses an Ad37 fiber or other D serotype fiber.

For exemplification, nucleic acid encoding GFP was incorporated into these vectors as a means to visualize their localization. Other genes, such as genes that encode therapeutic products, may be included in place of or in addition to GFP.

-25-

Plasmid pDV80 was electroporated into E1-2a S8 cells and stable lines were selected. The fiber-deleted vectors Ad5. β gal. Δ F and Ad5.GFP. Δ F were grown in cells in a resulting cell line, designated 705, to produce virions, which express the Ad37 fiber (Ad5. β gal. Δ F/37F and Ad5.GFP. Δ F/37F) and CsCl-
5 purified. These virions selectively transduce photoreceptor cells when injected intraocularly into the vitreous humor.

2. Packaging

Recombinant adenoviral vectors generally have at least a deletion in the first viral early gene region, referred to as E1, which includes the E1a and E1b
10 regions. Deletion of the viral E1 region renders the recombinant adenovirus defective for replication and incapable of producing infectious viral particles in subsequently-infected target cells. Thus, to generate E1-deleted adenovirus genome replication and to produce virus particles requires a system of complementation which provides the missing E1 gene product. E1
15 complementation is typically provided by a cell line expressing E1, such as the human embryonic kidney packaging cell line, i.e. an epithelial cell line, called 293. Cell line 293 contains the E1 region of adenovirus, which provides E1 gene region products to "support" the growth of E1-deleted virus in the cell line (see, e.g., Graham *et al.*, *J. Gen. Virol.* 36: 59-71, 1977). Additionally, cell lines
20 that may be usable for production of defective adenovirus having a portion of the adenovirus E4 region have been reported (WO 96/22378).

Multiply deficient adenoviral vectors and complementing cell lines have also been described (WO 95/34671, U.S. Patent No. 5,994,106).

Copending U.S. application Serial No. 09/482,682 (also filed as
25 International PCT application No. PCT/US00/00265, filed January 14, 2000)) provides packaging cell lines that support viral vectors with deletions of major portions of the viral genome, without the need for helper viruses and also provides cell lines and helper viruses for use with helper-dependent vectors. The packaging cell line has heterologous DNA stably integrated into the
30 chromosomes of the cellular genome. The heterologous DNA sequence encodes one or more adenovirus regulatory and/or structural polypeptides that complement the genes deleted or mutated in the adenovirus vector genome to

-26-

be replicated and packaged. The packaging cell line express, for example, one or more adenovirus structural proteins, polypeptides, or fragments thereof, such as penton base, hexon, fiber, polypeptide IIIa, polypeptide V, polypeptide VI, polypeptide VII, polypeptide VIII, and biologically active fragments thereof. The expression can be constitutive or under the control of a regulatable promoter. These cell lines are designed for expression of recombinant adenoviruses intended for delivery of therapeutic products.

Particular packaging cell lines complement viral vectors having a deletion or mutation of a DNA sequence encoding an adenovirus structural protein, regulatory polypeptides E1A and E1B, and/or one or more of the following regulatory proteins or polypeptides: E2A, E2B, E3, E4, L4, or fragments thereof.

The packaging cell lines are produced by introducing each DNA molecule into the cells and then into the genome via a separate complementing plasmid or plurality of DNA molecules encoding the complementing proteins can be introduced via a single complementing plasmid. Of interest herein, is a variation in which the complementing plasmid includes DNA encoding adenovirus fiber protein (or a chimeric or modified variant thereof), from Ad virus of subgroup D, such as Ad 37, polypeptide or fragment thereof.

For therapeutic applications, the delivery plasmid further includes a nucleotide sequence encoding a foreign polypeptide. Exemplary delivery plasmids include, but are not limited to, pDV44, p Δ E1B β -gal and p Δ E1sp1B. In a similar or analogous manner, therapeutic genes may be introduced.

The cell further includes a complementing plasmid encoding a fiber as contemplated herein; the plasmid or portion thereof is integrated into a chromosome(s) of the cellular genome of the cell.

In one embodiment, a composition comprises a cell containing first and second delivery plasmids wherein a first delivery plasmid comprises an adenovirus genome lacking a nucleotide sequence encoding fiber and incapable of directing the packaging of new viral particles in the absence of a second delivery plasmid, and a second delivery plasmid comprises an adenoviral genome capable of directing the packaging of new viral particles in the presence of the first delivery plasmid.

-27-

In a variation, the packaging cell line expresses fiber protein or chimeric variant thereof from an Ad virus of subgroup D, preferably Ad37, serotype or it can be any fiber protein but one that has been modified to include the portion of the Ad virus of subgroup D, such as Ad37, responsible for selective targeting to photoreceptors upon introduction into the vitreous humor of the eye of a mammal, preferably a human. The fiber protein can be further modified to include a non-native amino acid residue sequence that targets additional specific receptors. In all instances, the modification should not disrupt trimer formation or transport of fiber into the nucleus. In another variation, the non-native amino acid residue sequence alters the binding specificity of the fiber for a targeted cell type. The structural protein is fiber can include amino acid residue sequences from more than one adenovirus serotype. The nucleotide sequences encoding fiber protein or polypeptide need not be modified solely at one or both termini; fiber protein, may be modified "internally" as well as at the termini.

Additional nucleic acid fragments can encode polypeptides that are added to the fiber protein. In one variation, the non-native amino acid residue sequence is coupled to the carboxyl terminus of the fiber. In another, the non-native amino acid residue sequence further includes a linker sequence. Alternatively, the fiber protein further comprises a ligand coupled to the linker.

Suitable ligands include, but are not limited to, ligands that specifically or selectively bind to a cell surface receptor and ligands that can be used to couple other proteins or nucleic acid molecules. Typically, the packaging cell lines will contain nucleic acid encoding the fiber protein or modified protein stably integrated into a chromosome or chromosomes in the cellular genome.

The packaging cell line can be derived from a procaryotic cell line or from a eukaryotic cell line. While various embodiments suggest the use of mammalian cells, and more particularly, epithelial cell lines, a variety of other, non-epithelial cell lines are used in various embodiments. Thus, while various embodiments disclose the use of a cell line selected from among the 293, A549, W162, HeLa, Vero, 211, and 211A cell lines, and any other cell lines suitable for such use are likewise contemplated herein.

3. Components of the nucleic acid molecule included in the particle

-28-

A recombinant viral vector or therapeutic viral vector for use in the methods herein, typically includes a nucleic acid fragment that encodes a protein or polypeptide molecule, or a biologically active fragment thereof, or other regulatory sequence, that is intended for use in therapeutic applications.

- 5 The nucleic acid molecule to be packaged in the viral particle also may include an enhancer element and/or a promoter located 3' or 5' to and controlling the expression of the therapeutic product-encoding nucleic acid molecule if the product is a protein. Further, for purposes herein, the promoter and/or other transcriptional and translational regulatory sequences controlling
- 10 expression of the product is preferably one that is expressed specifically in the targeted cells, such as the a photoreceptor-specific promoter, such as a rhodopsin gene promoter.

- The nucleic acid molecule to be packaged in viral capsid includes at least 2 different operatively linked DNA segments. The DNA can be manipulated and
- 15 amplified by PCR as described herein and by using standard techniques, such as those described in *Molecular Cloning: A Laboratory Manual, 2nd Ed.*, Sambrook *et al.*, eds., Cold Spring Harbor, New York (1989). Typically, to produce such molecule, the sequence encoding the selected polypeptide and the promoter or enhancer are operatively linked to a DNA molecule capable of autonomous
- 20 replication in a cell either *in vivo* or *in vitro*. By operatively linking the enhancer element or promoter and nucleic acid molecule to the vector, the attached segments are replicated along with the vector sequences.

- Thus, the recombinant DNA molecule (rDNA) is a hybrid DNA molecule comprising at least 2 nucleotide sequences not normally found together in
- 25 nature. In various preferred embodiments, one of the sequences is a sequence encoding an Ad-derived polypeptide, protein, or fragment thereof. The nucleic acid molecule intended to be packaged is from about 20 base pairs to about 40,000 base pairs in length, preferably about 50 bp to about 38,000 bp in length. In various embodiments, the nucleic acid molecule is of sufficient length
- 30 to encode one or more adenovirus proteins or functional polypeptide portions thereof. Since individual Ad polypeptides vary in length from about 19 amino acid residues to about 967 amino acid residues, encoding nucleic acid molecules

-29-

from about 50 bp up to about 3000 bp, depending on the number and size of individual polypeptide-encoding sequences that are "replaced" in the viral vectors by therapeutic product-encoding nucleic acid molecules.

Preferably the molecule includes an adenovirus tripartite leader (TPL)

- 5 nucleic acid sequence operatively linked to an intron containing RNA processing signals (such as for example, splice donor or splice acceptor sites) suitable for expression in the packaging cell line. Most preferably the intron contains a splice donor site and a splice acceptor site. Alternatively, the TPL nucleotide sequence may not comprise an intron. The intron includes any sequence of
- 10 nucleotides that function in the packaging cell line to provide RNA processing signals, including splicing signals. Introns have been well characterized from a large number of structural genes, and include but are not limited to a native intron 1 from adenovirus, such as Ad5's TPL intron 1; others include the SV40 VP intron; the rabbit beta-globin intron, and synthetic intron constructs (see,
- 15 *e.g.*, Petitclerc *et al.* (1995) *J. Biootechnol.*, 40:169; and Choi *et al.* (1991) *Mol. Cell. Biol.*, 11:3070).

The nucleic acid molecule encoding the TPL includes either (a) first and second TPL exons or (b) first, second and third TPL exons, where each TPL exon in the sequence is selected from among the complete TPL exon 1, partial TPL

20 exon 1, complete TPL exon 2 and complete TPL exon 3. A complete exon is one which contains the complete nucleic acid sequence based on the sequence found in the wild type viral genome. Preferably the TPL exons are from Ad2, Ad3, Ad5, Ad7 and the like, however, they may come from any Ad serotype, as described herein. A preferred partial TPL exon 1 is described in the Examples.

- 25 The use of a TPL with a partial exon 1 has been reported (International PCT application No. WO 98/13499).

- The intron and the TPL exons can be operatively linked in a variety of configurations to provide a functional TPL nucleotide sequence. An intron may not be a part of the construct. For example, the intron can be positioned
- 30 between any of TPL exons 1, 2 or 3, and the exons can be in any order of first and second, or first/second/third. The intron can also be placed preceding the first TPL exon or following the last TPL exon. In a preferred embodiment,

-30-

complete TPL exon 1 is operatively linked to complete TPL exon 2 operatively linked to complete TPL exon 3. In a preferred variation, adenovirus TPL intron 1 is positioned between complete TPL exon 1 and complete TPL exon 2. It may also be possible to use analogous translational regulators from other viral systems such as rabiesvirus.

5 A preferred "complete" TPL nucleic acid molecule containing complete TPL exons 1, 2 and 3 with adenovirus intron 1 inserted between exons 1 and 2 has a nucleotide sequence shown in SEQ ID NO: 32. A preferred "partial" TPL nucleic acid molecule containing partial TPL exon 1 and complete TPL exons 2
10 and 3 in that order has a nucleotide sequence shown in SEQ ID NO: 26. The construction of these preferred TPL nucleotide sequences is described in the Examples.

Thus, preferred expression cassettes and complementing plasmids for expressing adenovirus structural genes, particularly fiber protein, contain an
15 adenovirus TPL nucleotide sequence as described herein.

4. Complementing Plasmids

Also contemplated are the use of nucleic acid molecules, typically in the form of DNA plasmid vectors, which are capable of expression of an adenovirus structural protein or regulatory protein. Because these expression plasmids are
20 used to complement the defective genes of a recombinant adenovirus vector genome, the plasmids are referred to as complementing or complementation plasmids.

The complementing plasmid contains an expression cassette, a nucleotide sequence capable of expressing a protein product encoded by the nucleic acid
25 molecule. Expression cassettes typically contain a promoter and a structural gene operatively linked to the promoter. The complementing plasmid can further include a sequence of nucleotides encoding TPL nucleotide to enhance expression of the structural gene product when used in the context of adenovirus genome replication and packaging.

30 A complementing plasmid can include a promoter operatively linked to a sequence of nucleotides encoding an adenovirus structural polypeptide, such as, but are not limited to, penton base; hexon; fiber; polypeptide IIIa; polypeptide V;

-31-

polypeptide VI; polypeptide VII; polypeptide VIII; and biologically active fragments thereof. In another variation, a complementing plasmid may also include a sequence of nucleotides encoding a first adenovirus regulatory polypeptide, a second regulatory polypeptide, and/or a third regulatory polypeptide, and any combination of the foregoing.

5 Plasmid pDV80 is a preferred plasmid herein. Other plasmids constructed in an analogous manner to encode modified fiber proteins and chimeric fiber proteins are also contemplated herein.

5. Nucleic Acid Molecule Synthesis

10 A nucleic acid molecule comprising synthetic oligonucleotides can be prepared using any suitable method, such as the phosphotriester or phosphodiester methods (see, *e.g.*, Narang (1979) *et al.*, *Meth. Enzymol.*, 68:90; U.S. Patent No. 4,356,270; and Brown *et al.*, (1979) *Meth. Enzymol.*, 68:109). For oligonucleotides, the synthesis of the family members can be

15 conducted simultaneously in a single reaction vessel, or can be synthesized independently and later admixed in preselected molar ratios. For simultaneous synthesis, the nucleotide residues that are conserved at preselected positions of the sequence of the family member can be introduced in a chemical synthesis protocol simultaneously to the variants by the addition of a single preselected

20 nucleotide precursor to the solid phase oligonucleotide reaction admixture when that position number of the oligonucleotide is being chemically added to the growing oligonucleotide polymer. The addition of nucleotide residues to those positions in the sequence that vary can be introduced simultaneously by the addition of amounts, preferably equimolar amounts, of multiple preselected

25 nucleotide precursors to the solid phase oligonucleotide reaction admixture during chemical synthesis. For example, where all four possible natural nucleotides (A,T,G and C) are to be added at a preselected position, their precursors are added to the oligonucleotide synthesis reaction at that step to simultaneously form four variants (see, *e.g.*, Ausubel *et al.* (*Current Protocols in*

30 *Molecular Biology*, Suppl. 8. p.2.11.7, John Wiley & Sons, Inc., New York, 1991).

-32-

Nucleotide bases other than the common four nucleotides (A,T,G or C), or the RNA equivalent nucleotide uracil (U), can also be used. For example, it is well known that inosine (I) is capable of hybridizing with A, T and G, but not C. Examples of other useful nucleotide analogs are known in the art and may be found referred to in 37 C.F.R. §1.822.

Thus, where all four common nucleotides are to occupy a single position of a family of oligonucleotides, that is, where the preselected nucleotide sequence is designed to contain oligonucleotides that can hybridize to four sequences that vary at one position, several different oligonucleotide structures are contemplated. The composition can contain four members, where a preselected position contains A,T,G or C. Alternatively, a composition can contain two nucleotide sequence members, where a preselected position contains I or C, and has the capacity to hybridize at that position to all four possible common nucleotides. Finally, other nucleotides may be included at the preselected position that have the capacity to hybridize in a non-destabilizing manner with more than one of the common nucleotides in a manner similar to inosine.

Similarly, larger nucleic acid molecules can be constructed in synthetic oligonucleotide pieces, and assembled by complementary hybridization and ligation, as is well known.

D. Adenovirus Expression Vector Systems

The adenovirus vector genome that is encapsulated in the virus particle and that expresses exogenous genes in a gene therapy setting is a key component of the system. Thus, the components of a recombinant adenovirus vector genome include the ability to express selected adenovirus structural genes, to express a desired exogenous protein, and to contain sufficient replication and packaging signals that the genome is packaged into a gene delivery vector particle. The preferred replication signal is an adenovirus inverted terminal repeat containing an adenovirus origin of replication, as is well known and described herein.

-33-

Although adenovirus include many proteins, not all adenovirus proteins are required for assembly of a recombinant adenovirus particle (vector). Thus, deletion of the appropriate genes from a recombinant Ad vector permits accommodation of even larger "foreign" DNA segments.

- 5 A preferred recombinant adenovirus vector genome is "helper independent" so that genome can replicate and be packaged without the help of a second, complementing helper virus. Complementation is provided by a packaging cell.

- 10 In a preferred embodiment, the adenovirus vector genome does not encode a functional adenovirus fiber protein. A non-functional fiber gene refers to a deletion, mutation or other modification to the adenovirus fiber gene such that the gene does not express any or insufficient adenovirus fiber protein to package a fiber-containing adenovirus particle without complementation of the fiber gene by a complementing plasmid or packaging cell line. Such a genome is
- 15 referred to as a "fiberless" genome, not to be confused with a fiberless particle. Alternatively, a fiber protein may be encoded but is insufficiently expressed to result in a fiber containing particle.

- 20 Thus, contemplated for use are helper-independent fiberless recombinant adenovirus vector genomes that include genes that (a) express all adenovirus structural gene products but express insufficient adenovirus fiber protein to package a fiber-containing adenovirus particle without complementation of said fiber gene, (b) express an exogenous protein, and (c) contain an adenovirus packaging signal and inverted terminal repeats containing adenovirus origin of replication.

- 25 The adenovirus vector genome is propagated in the laboratory in the form of rDNA plasmids containing the genome, and upon introduction into an appropriate host, the viral genetic elements provide for viral genome replication and packaging rather than plasmid-based propagation. Exemplary methods for preparing an Ad-vector genome are described in the Examples.

- 30 A vector herein includes a nucleic acid (preferably DNA) molecule capable of autonomous replication in a cell and to which a DNA segment, *e.g.*, a gene or polynucleotide, can be operatively linked to bring about replication of the

-34-

attached segment. For purposes herein, one of the nucleotide segments to be operatively linked to vector sequences encodes at least a portion of a therapeutic nucleic acid molecule. As noted above, therapeutic nucleic acid molecules include those encoding proteins and also those that encode regulatory factors that can lead to expression or inhibition or alteration of expression of a gene product in a targeted cell.

1. Nucleic Acid Gene Expression Cassettes

In various embodiments, a peptide-coding sequence of the therapeutic gene is inserted into an expression vector and expressed; however, it is also feasible to construct an expression vector which also includes some non-coding sequences as well. Preferably, however, non-coding sequences are excluded. Alternatively, a nucleotide sequence for a soluble form of a polypeptide may be utilized. Another preferred therapeutic viral vector includes a nucleotide sequence encoding at least a portion of a therapeutic nucleotide sequence operatively linked to the expression vector for expression of the coding sequence in the therapeutic nucleotide sequence.

The choice of viral vector into which a therapeutic nucleic acid molecule is operatively linked depends directly, as is well known in the art, on the functional properties desired, *e.g.*, vector replication and protein expression, and the host cell to be transformed -- these being limitations inherent in the art of constructing recombinant DNA molecules. Although certain adenovirus serotypes are recited herein in the form of specific examples, it should be understood that the use of *any* adenovirus serotype, including hybrids and derivatives thereof are contemplated.

A translatable nucleotide sequence is a linear series of nucleotides that provide an uninterrupted series of at least 8 codons that encode a polypeptide in one reading frame. Preferably, the nucleotide sequence is a DNA sequence. The vector itself may be of any suitable type, such as a viral vector (RNA or DNA), naked straight-chain or circular DNA, or a vesicle or envelope containing the nucleic acid material and any polypeptides that are to be inserted into the cell.

-35-

2. Promoters

As noted elsewhere herein, an expression nucleic acid in an Ad-derived vector may also include a promoter, particularly a tissue or cell specific promoter, preferably one expressed in ocular cells, particularly photoreceptors.

- 5 Promoters contemplated for use herein include regulatable (inducible) as well as constitutive promoters, which may be used, either on separate vectors or on the same vector. Some useful regulatable promoters are those of the CREB-regulated gene family and include inhibin, gonadotropin, cytochrome c, glucagon, and the like. (See, *e.g.*, International PCT application No. WO 10 96/14061). Preferably the promoter selected is from a photoreceptor-specific gene, such as a rhodopsin gene or gene that encodes a protein that regulates rhodopsin expression.

E. Formulation and administration

- Compositions containing therapeutically effective concentrations of 15 recombinant adenovirus delivery vectors are provided. These are for delivery of therapeutic gene products to cells, particularly cells express a particular 50 kDa receptor or other receptor with which the vectors interact. These cells include cells of the eye and genital tract. Of particular interest are photoreceptor cells of the eye. Administration is effected by any means through which contacting with 20 the photoreceptors is effected. Preferable modes of administration include, but are not limited to, subretinal injection, particularly intravitreal injection, to provide access to photoreceptor cells.

- The recombinant viral compositions may also be formulated for implantation into the anterior or posterior chamber of the eye, preferably the 25 vitreous cavity, in sustained released formulations, such as those adsorbed to biodegradable supports, including collagen sponges, or in liposomes. Sustained release formulations may be formulated for multiple dosage administration, so that during a selected period of time, such as a month or up to about a year, several dosages are administered. Thus, for example, liposomes may be 30 prepared such that a total of about two to up to about five or more times the single dosage is administered in one injection.

-36-

The vectors are formulated in an ophthalmologically acceptable carrier for intraocular, preferably intravitreal, administration in a volume of between about 0.05 ml and 0.150 ml, preferably about 0.05 and 0.100 ml.

The composition can be provided in a sealed sterile vial containing an amount of a compound of formula I, that upon intraocular administration will deliver a sufficient amount of viral particles to the photoreceptors in a volume of about 50 to 150 μ l, containing at least about 10^7 , more preferably at least about 10^8 plaque forming units in such volume. Typically, the vials will, thus, contain about 0.150 ml of the composition.

To prepare compositions the viral particles are dialyzed into a suitable ophthalmologically acceptable carrier or viral particles, for example, may be concentrated and/or mixed therewith. The resulting mixture may be a solution, suspension or emulsion. In addition, the viral particles may be formulated as the sole pharmaceutically active ingredient in the composition or may be combined with other active agents for the particular disorder treated.

For administration by intraocular injection or via eyedrops, suitable carriers include, but are not limited to, physiological saline, phosphate buffered saline (PBS), balanced salt solution (BSS), lactate Ringers solution, and solutions containing thickening and solubilizing agents, such as glucose, polyethylene glycol, and polypropylene glycol and mixtures thereof. Liposomal suspensions may also be suitable as pharmaceutically acceptable carriers. These may be prepared according to methods known to those skilled in the art. Suitable ophthalmologically acceptable carriers are known. Solutions or mixtures intended for ophthalmic use may be formulated as 0.01% - 10% isotonic solutions, pH about 5-7, with appropriate salts [see, e.g., U.S. Patent No. 5,116,868, which describes typical compositions of ophthalmic irrigation solutions and solutions for local application]. Such solutions, which have a pH adjusted to about 7.4, contain, for example, 90-100 mM sodium chloride, 4-6 mM dibasic potassium phosphate, 4-6 mM dibasic sodium phosphate, 8-12 mM sodium citrate, 0.5-1.5 mM magnesium chloride, 1.5-2.5 mM calcium chloride, 15-25 mM sodium acetate, 10-20 mM D.L.-sodium β -hydroxybutyrate and 5-5.5 mM glucose.

-37-

The compositions may be prepared with carriers that protect them from rapid elimination from the body, such as time release formulations or coatings. Such carriers include controlled release formulations, such as, but not limited to, microencapsulated delivery systems, and biodegradable, biocompatible polymers, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, polyorthoesters, polylactic acid and other types of implants that may be placed directly into the anterior or posterior chamber or vitreous cavity of the eye. The compositions may also be administered in pellets, such as Elvax pellets (ethylene-vinyl acetate copolymer resin).

- 10 Liposomal suspensions, including tissue-targeted liposomes, may also be suitable as pharmaceutically acceptable carriers. For example, liposome formulations may be prepared by methods known to those of skill in the art [see, e.g., Kimm et al. (1983) Bioch. Bioph. Acta 728:339-398; Assil et al. (1987) Arch Ophthalmol. 105:400; and U.S. Patent No. 4,522,811]. The viral particles
- 15 may be encapsulated into the aqueous phase of liposome systems.

- The active materials can also be mixed with other active materials, that do not impair the desired action, or with materials that supplement the desired action or have other action, including viscoelastic materials, such as hyaluronic acid, which is sold under the trademark HEALON, which is a solution of a high
- 20 molecular weight (MW) of about 3 millions fraction of sodium hyaluronate [manufactured by Pharmacia, Inc; see, e.g., U.S. Patent Nos. 5,292,362, 5,282,851, 5,273,056, 5,229,127, 4,517,295 and 4,328,803], VISCOAT [fluorine-containing (meth)acrylates, such as, 1H,1H,2H,2H-hepta-decafluorodecylmethacrylate; see, e.g., U.S. Patent Nos. 5,278,126, 5,273,751
- 25 and 5,214,080; commercially available from Alcon Surgical, Inc.], ORCOLON [see, e.g., U.S. Patent No. 5,273,056; commercially available from Optical Radiation Corporation], methylcellulose, methyl hyaluronate, polyacrylamide and polymethacrylamide [see, e.g., U.S. Patent No. 5,273,751]. The viscoelastic
- 30 materials are present generally in amounts ranging from about 0.5 to 5.0%, preferably 1 to 3% by weight of the conjugate material and serve to coat and protect the treated tissues. The compositions may also include a dye, such as

-38-

methylene blue or other inert dye, so that the composition can be seen when injected into the eye. Additional active agents may be included.

The compositions can be enclosed in ampules, disposable syringes or multiple or single dose vials made of glass, plastic or other suitable material.

- 5 Such enclosed compositions can be provided in kits. In particular, kits containing vials, ampules or other containers, preferably disposable vials with sufficient amount of the composition to deliver about 0.100 ml thereof, and disposable needles, preferably self sealing 25-30 gauge needles, are provided herein.
- 10 Finally, the compounds may be packaged as articles of manufacture containing packaging material, typically a vial, an ophthalmologically acceptable composition containing the viral particles and a label that indicates the therapeutic use of the composition.

- Also provided are kits for practice of the methods herein. The kits
- 15 contain one or more containers, such as sealed vials, with sufficient composition for single dosage administration, and one or more needles, such as self sealing 25-33 gauge needles, preferably 33 gauge or smaller needles, precisely calibrated syringes or other precisely calibrated delivery device, suitable for intravitreal injection.

- 20 Administration of the composition is preferably by intraocular injection, although other modes of administration may be effective, if the sufficient amount of the compound achieves contact with the vitreous cavity. Intraocular injection may be effected by intravitreal injection, aqueous humor injection or injection into the external layers of the eye, such as subconjunctival injection or
- 25 subtenon injection, or by topical application to the cornea, if a penetrating formulation is used.

Administration

- The compositions containing the compounds are administered intraocularly or by other means, such as topically in the form of penetrating
- 30 eyedrops, whereby contact of the recombinant vectors with the aqueous humor is effected. Intraocular administration may be effected by intravitreal injection, aqueous humor injection, injection into the external layers of the eye, such as

-39-

subconjunctival injection or subtenon injection, preferably in free form, but, alternatively, in liposomes or other sustained drug delivery device. Administration is preferably by intravitreal injection, preferably through self sealing 25-30 gauge needles or other suitably calibrated delivery device.

- 5 Injection into the eye may be through the pars plana via the self-sealing needle.

It is further understood that, for any particular subject, specific dosage regimens should be adjusted over time according to the individual need and the professional judgment of the person administering or supervising the administration of the recombinant viruses, and that the concentration ranges set forth herein are exemplary only and are not intended to limit the scope or practice of the claimed methods

F. Diseases, Disorders and therapeutic products

1. Disease and disorders

Retinitis pigmentosa

- 15 Methods for specifically or selectively targeting recombinant adenovirus vectors for delivery of gene products, particularly therapeutic products are provided herein. These methods are particularly suitable for targeting cells that express receptors that are selectively recognized by Ad virus of subgroup D viruses, particularly Ad37. It is shown herein that these viruses selectively
- 20 recognize receptors on cells, such as conjunctival cells and photoreceptors, that are not recognized by other adenoviruses. Hence, methods for targeting to these cell types by providing vectors that are packaged in viral particles that contain a sufficient portion of a fiber protein from one of these Ad serotypes to bind to these receptors. These methods are useful for targeting to
- 25 photoreceptors and for treating ocular disorders, including, but are not limited to, inherited and acquired retinal, neovascular degenerative diseases (see table below).

- It is estimated that 1 in 3,500 individuals in the United States suffer from one of the pigmented retinopathies. This group of retinal diseases, commonly
- 30 called retinitis pigmentosa, is characterized by progressive loss of peripheral and night vision. Patients may be affected at almost any age and it is not uncommon to experience symptoms in early childhood in certain inherited forms.

-40-

It has been shown that there are a variety of mutations in genes expressed in the photoreceptors, including genes in the rhodopsin gene and pathway that appear to be responsible for these diseases. In addition to mutations in rhodopsin, changes in the retinal pigmented epithelial (RPE) cells, also undergo
5 degenerative changes and can form clumps of pigment that give rise to the characteristic pigmentary changes seen in patients with RP.

Angiogenesis and ocular diseases and disorders

The vast majority of diseases that cause catastrophic loss of vision do so as a result of ocular neovascularization; age related macular degeneration
10 (ARMD) affects 12-15 million American over the age of 65 and causes visual loss in 10-15% of them as a direct effect of choroidal (sub-retinal) neovascularization. The leading cause of visual loss for Americans under the age of 65 is diabetes; 16 million individuals in the United States are diabetic and 40,000 per year suffer from ocular complications of the disease, which often are
15 a result of retinal neovascularization. Laser photocoagulation has been effective in preventing severe visual loss in subgroups of high risk diabetic patients, but the overall 10 year incidence of retinopathy remains essentially unchanged. For patients with choroidal neovascularization due to ARMD or inflammatory eye disease, such as ocular histoplasmosis, photocoagulation, with few exceptions,
20 is ineffective in preventing visual loss. While recently developed, non-destructive photodynamic therapies hold promise for temporarily reducing individual loss in patients with previously untreatable choroidal neovascularization, only 61.4% of patients treated every 3-4 months had improved or stabilized vision compared to 45.9% of the placebo-treated group.
25 In the normal adult, angiogenesis is tightly regulated and limited to wound healing, pregnancy and uterine cycling. Angiogenesis is turned on by specific angiogenic molecules such as basic and acidic fibroblast growth factor (FGF), vascular endothelial growth factor (VEGF), angiogenin, transforming growth factor (TGF), tumor necrosis factor- α (TNF- α) and platelet derived growth factor
30 (PDGF). Angiogenesis can be suppressed by inhibitory molecules such as interferon- α , thrombospondin-1, angiostatin and endostatin. It is the balance of these naturally occurring stimulators and inhibitors that controls the normally

-41-

quiescent capillary vasculature. When this balance is upset, as in certain disease states, capillary endothelial cells are induced to proliferate, migrate and ultimately differentiate.

Angiogenesis plays a central role in a variety of diseases, including, but are not limited to, cancer and ocular neovascularization. Sustained growth and metastasis of a variety of tumors has also been shown to be dependent on the growth of new host blood vessels into the tumor in response to tumor derived angiogenic factors. Proliferation of new blood vessels in response to a variety of stimuli occurs as the dominant finding in the majority of eye diseases that blind, such as, but are not limited to, proliferative diabetic retinopathy (PDR), ARMD, rubeotic glaucoma, interstitial keratitis and retinopathy of prematurity. In these diseases, tissue damage can stimulate release of angiogenic factors resulting in capillary proliferation. VEGF plays a dominant role in iris neovascularization and neovascular retinopathies. While reports clearly show a correlation between intraocular VEGF levels and ischemic retinopathic ocular neovascularization, FGF likely plays a role. Basic and acidic FGF are known to be present in the normal adult retina, even though detectable levels are not consistently correlated with neovascularization. This may be largely due to the fact that FGF binds very tightly to charged components of the extracellular matrix and may not be readily available in a freely diffusible form that would be detected by standard assays of intraocular fluids.

A final common pathway in the angiogenic response involves integrin-mediated information exchange between a proliferating vascular endothelial cell and the extracellular matrix. This class of adhesion receptors, called integrins, are expressed as heterodimers having an α and β subunit on all cells. One such integrin, $\alpha_v\beta_3$, is the most promiscuous member of this family and allows endothelial cells to interact with a wide variety of extracellular matrix components. Peptide and antibody antagonists of this integrin inhibit angiogenesis by selectively inducing apoptosis of the proliferating vascular endothelial cells. Two cytokine-dependent pathways of angiogenesis exist and may be defined by their dependency on distinct vascular cell integrins, $\alpha_v\beta_3$ and $\alpha_v\beta_5$. Specifically, basic FGF- and VEGF-induced angiogenesis depend on integrin

-42-

$\alpha_v\beta_3$ and $\alpha_v\beta_5$, respectively, since antibody antagonists of each integrin selectively block one of these angiogenic pathways in the rabbit corneal and chick chorioallantoic membrane (CAM) models. Peptide antagonists that block all α_v integrins inhibit FGF- and VEGF-stimulated angiogenesis. While normal human ocular blood vessels do not display either integrin, $\alpha_v\beta_3$ and $\alpha_v\beta_5$ integrins are selectively displayed on blood vessels in tissues from patients with active neovascular eye disease. While only $\alpha_v\beta_3$ was consistently observed in tissue from patients with ARMD, $\alpha_v\beta_3$ and $\alpha_v\beta_5$ were present in tissues from patients with PDR. Systemically administered peptide antagonists of integrins blocked new blood vessel formation in a mouse model of retinal vasculogenesis.

In addition to adhesion events described above, cell migration through the extracellular matrix also depends on proteolysis. Matrix metalloproteinases are a family of zinc-requiring matrix-degrading enzymes that include the collagenases, gelatinases and stromelysins, all of which have been implicated in invasive cell behavior. Invasive cell processes such as tumor metastasis and angiogenesis have been found to be associated with the expression of integrins and MMP-2, MMP-2 are all found throughout the eye where they may interact to maintain a quiescent vasculature until the balance is upset, resulting in pathological angiogenesis. A non-catalytic C-terminal hemopexin-like domain of MMP-2 (PEX) can block cell surface collagenolytic activity and inhibit angiogenesis in the CAM model by preventing localization of MMP-2 to the surface of invasive cells through interaction with the integrin $\alpha_v\beta_3$.

Hence, anti-angiogenic agents have a role in treating retinal degeneration to prevent the damaging effects of these trophic and growth factors. Angiogenic agents, also have a role in promoting desirable vascularization to retard retinal degeneration by enhancing blood flow to cells.

Members of adenovirus subgroup D, Ad8, 19A, and 37, are infectious agents that cause particularly severe cases of epidemic keratoconjunctivitis (EKC) (Arnberg *et al.* (1998) *Virology* 227:239-244; Curtis *et al.* (1998) *J. Med. Microbiol.* 47:91-94; Ritterband *et al.* (1998) *Rev. Med. Virol.* 8:187-201; and Takeuchi *et al.* (1999) *J. Clin. Microbiol.* 37:3392-3394). There is no effective treatment for this debilitating and contagious disease and EKC

-43-

continues to be a problem in ophthalmology clinics worldwide (Curtis *et al.* (1998) *J.Med.Microbiol.* 47:91-94, Lukashok *et al.* (1998) *Curr.Clin.Top.Infec.Dis.* 18:286-304). Hence the vectors herein may be used for treating the disease.

5

Table 3
Candidate targets for ocular disease therapy

CANDIDATE TARGETS FOR OCULAR DISEASE THERAPY	
Disease	Candidate target(s)
Retinitis pigmentosa	Rhodopsin gene, and genes that regulate expression thereof <i>rds/peripherin</i>
10 Stargardt's disease	<i>rim</i> protein (ARC protein)
Choroideremia	<i>rab</i> geranylgeranyl transferase CHM, TCD, CHML *
Gyrate Atrophy	ornithine aminotransferase
15 Macular dystrophy	<i>rds/peripherin</i>

15

* see, "MSR6-yeast homologue of the choroideraemia gene," *Nature Genetics* 3: 193-4 (1993)

TABLE 4

Other Diseases	
Exudative Choroidal Diseases	
20	ICSC, fluorescein angiogram
	ICSC with large serous detachment of RPE (retinal pigmented epithelium)
	ICSC with bullous retinal detachment
	Macular drusen, exudative, confluent
25	Drusen, sub-RPE choroidal neovascularization
	Drusen, notched serous detachment of RPE

Other Diseases	
	Drusen, notched serous and hemorrhagic detachment of RPE
	Drusen, serous and hemorrhagic detachment of RPE and retina
	Drusen, organized RPE detachment causing bullous retinal detachment
	Drusen, geographic atrophy of RPE
5	Drusen, exudative and cuticular, vitelliform macular detachment
	Drusen, cuticular, large vitelliform macular detachment
	North Carolina dystrophy with macular staphyloma
	North Carolina dystrophy with macular staphyloma
	Angioid streaks, pseudoxanthoma elasticum (PXE), CNVM
10	Angioid streaks, PXE, large notched retinal detachment
	Myopic degeneration, Foerster-Fuchs spot
	Presumed ocular histoplasmosis syndrome (POHS)
	Submacular bacterial abscess
	<i>Toxocara canis</i> , subretinal granuloma
15	Serpiginous (geographic) choroiditis
	Posterior scleritis
	Harada's disease
	Posterior sympathetic uveitis
	Benign reactive lymphoid hyperplasia of uveal tract
20	Choroidal ruptures and CNVM
	Cavernous hemangioma of choroid
	Choroidal osteoma
	Choroidal nevus, serous macular detachment
	Choroidal nevus with CNVM
25	Diffuse sclerochoroidal melanocytic nevus
	Choroidal melanoma with serous detachment of RPE
	Metastatic lung carcinoma to choroid
	Sub-RPE reticulum cell sarcoma

-45-

Other Diseases		
	RPE tear, idiopathic choroidal neovascularization	
Heredodystrophic Disorders Affecting RPE & Retina		
	Best's vitelliform macular dystrophy	
	Best's vitelliform macular dystrophy with CNVM	
5	Best's vitelliform macular dystrophy, multiple lesions	
	Adult-onset vitelliform foveomacular dystrophy	
	Pattern dystrophy simulating fundus flavimaculatus	
	Stargardt's disease (fundus flavimaculatus)	
	Asteroid macular dystrophy	
10	Sjögren-Larssen syndrome	
	Oguchi's disease, light-adapted state	
	Oguchi's disease, dark-adapted state	
	Fundus albipunctatus	
	Retinitis pigmentosa, cystoid macular edema	
15	Crystalline tapetoretinal dystrophy	
	Choroideremia	
	Goldmann-Favre syndrome	
	Sex-linked juvenile retinoschisis	
	Perivenous retinitis pigmentosa	
20	Retinal Vascular Disorders	
	Retinal arteriovenous aneurysm	
	Central retinal artery occlusion	
	Cilioretinal artery obstruction	
	Ischemic retinopathy in systemic lupus erythematosus	
25	Ischemic retinopathy in scleroderma	
	Hemorrhagic detachment of internal limiting membrane, hypertensive retinopathy	
	Acquired retinal arterial macroaneurysm	

Other Diseases	
	Cystoid macular edema, aphakic
	Cystoid macular edema, nicotinic acid maculopathy
	Congenital retinal telangiectasis
	Acquired bilateral juxtafoveal telangiectasis
5	Acquired bilateral juxtafoveal obliterative telangiectasis
	Diabetic optic neuropathy
	X-ray radiation exudative retinopathy
	Sickle cell SC disease, macular hemorrhage
	Retinal arterial aneurysms, arteritis, neuroretinitis
10	Branch retinal vein obstruction (BRVO)
	BRVO, exudative maculopathy
	BRVO, optic disc new vessels, photocoagulation
	Waldenström's macroglobulinemia
Inflammatory Diseases of the Retina and Choroid	
15	Luetic retinal vasculitis
	Focal <i>Candida</i> retinal abscess
	Toxoplasmosis, atrophic chorioretinal scar
	Toxoplasmosis retinitis and macular detachment
	Toxoplasmosis scar, CNVM, macular detachment
20	Diffuse unilateral subacute neuroretinitis, small worm
	Diffuse unilateral subacute neuroretinitis, large worm
	Cytomegalic inclusion disease, papillitis
	Acute posterior multifocal placoid pigment epitheliopathy
	Acute macular neuroretinitis
25	Sarcoid retinitis
	Sarcoid papillitis
	Behcet's disease
	Vitiliginous (bird-shot) chorioretinitis

	Other Diseases
	Multifocal choroiditis and panveitis (pseudo-POHS)
	Retinal and Pigment Epithelial Hamartomas
	Congenital grouped albinotic RPE spots
	Congenital hyperplasia of RPE
5	Combined RPE and retinal hamartoma, juxtapapillary
	Combined RPE and retinal hamartoma, peripheral
	Cystic astrocytoma, juxtapapillary
	Astrocytoma, macula
	Astrocytoma, juxtapapillary
10	Cavernous hemangioma of retina
	Juxtapapillary sessile retinal capillary hemangioma
	Juxtapapillary endophytic retinal capillary hemangioma
	Other Tumors of the Choroid
	Choroidal metastasis
15	Choroidal osteoma
	Choroidal hemangioma
	Miscellaneous uveal tumors
	Intraocular Lymphoid Tumors
	The leukemias and lymphomas
20	Tumors of the Vitreous
	Non-Hodgkins ("reticulum cell") lymphoma
	Tumor involvement of the vitreous cavity
	Macular Disease
	Age-related macular degeneration -- atrophic form
25	Exudative age-related macular degeneration
	Choroidal neovascular membrane in degenerative myopia
	Central serous retinopathy
	Macular hole

	Other Diseases
	Macular dystrophies
	Retinal Vascular Disease
	Etiologic mechanisms in diabetic retinopathy
	Background diabetic retinopathy
5	Proliferative diabetic retinopathy
	Retinal arterial obstructive disease
	Central retinal vein occlusion
	Retinal branch vein occlusion
	Pregnancy and retinal disease
10	Pregnancy-induced hypertension
	Hypertension
	The rheumatic disease
	Parafoveal telangiectasis
	Coats disease
15	Disseminated intravascular systemic coagulopathy and related vasculopathies
	Hemoglobinopathies
	Retinopathy of prematurity
	Acquired retinal macroaneurysms
	Eales disease
20	Radiation retinopathy
	The ocular ischemic syndrome
	Inflammatory Disease
	Ocular toxoplasmosis
	Ocular toxocariasis
25	Ocular cysticercosis
	Cytomegalovirus infections of the retina
	Retinal and ophthalmologic manifestations of AIDS
	Acute retinal necrosis syndrome

-49-

Other Diseases	
	Endogenous fungal infections of the retina and choroid
	Pars planitis
	Syphilis and tuberculosis
	Diffuse unilateral subacute neuroretinitis
5	Scleritis
	Birdshot retinochoroidopathy
	Punctate inner choroidopathy
	Sarcoidosis
	Acute multifocal placoid pigment epitheliopathy
10	Geographic helicoid peripapillary choroidopathy (GHPC): serpiginous choroiditis
	Sympathetic ophthalmia
	Vogt-Koyanagi-Harada syndrome (uveomeningitic syndrome)
	Ciliochoroidal (uveal) effusion

15 Reproduced from: Stereoscopic Atlas of Ocular Diseases Diagnosis and Treatment, 2nd Edition, J. Donald O. Gass, Vol. 1 & 2, C.V. Mosley Co. (1987); and Retina Vol. II, Editor, Stephen J. Ryan, Medical Retina, C.V. Mosley Co. (1989).

20 2. Therapeutic products

Therapeutic products include but are not limited to, wild-type genes that are defective in ocular disorders, such as rhodopsin, or fragments thereof sufficient to correct the genetic defect, trophic factors, including growth factors, inhibitors and agonists of trophic factors, anti-apoptosis factors and other products described herein or known to those of skill in the art to be useful for treatment of disorders of the eye or that can be treated by a product expressed by a photoreceptor.

OCULAR GENE THERAPY STRATEGIES		
GENERAL DISEASE	EXAMPLES	STRATEGY

-50-

OCULAR GENE THERAPY STRATEGIES		
Hereditary retinal and macular degeneration	<ul style="list-style-type: none"> • Retinitis pigmentosa • Stargardt's disease • Other macular dystrophies 	Growth factors (e.g., GDNF) anti-apoptotic factors (e.g., bcl2 gene) Stargardt Disease Gene (ABCR) [†]
Neovascular	<ul style="list-style-type: none"> • Diabetes • Choroidal neovascularization 	Anti-angiogenesis factors
Anti-tumor	Retinoblastoma	Antiproliferant
5 Glaucoma	Nerve fiber layer atrophy	Neuroprotective agent

[†] See Allikmets *et al.* (1997) *Science* 277:1805-1807.

For example, for treatment of retinitis pigmentosa the adenovirus vector can deliver a wild-type rhodopsin gene or a growth factor or trophic factor, such as ciliary neurotrophic factor CNTF; for treatment of Stargardt's disease, the
 10 vector can deliver a wild type ABCR (also called STGD1) or a growth factor or anti-angiogenic agent; for diabetic retinopathies, retinal vascularization the vector can deliver growth factors, such as a TGF (TGF β), to prevent degeneration.

15 The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLE 1

Preparation of Adenovirus Packaging Cell Lines

Cell lines that are commonly used for growing adenovirus are useful as
 20 host cells for the preparation of adenovirus packaging cell lines. Preferred cells include 293 cells, an adenovirus-transformed human embryonic kidney cell line obtained from the ATCC, having Accession Number CRL 1573; HeLa, a human epithelial carcinoma cell line (ATCC Accession Number CCL-2); A549, a human lung carcinoma cell line (ATCC Accession Number CCL 1889); and other
 25 epithelial-derived cell lines. As a result of the adenovirus transformation, the

-51-

293 cells contain the E1 early region regulatory gene. All cells were maintained in complete DMEM + 10% fetal calf serum unless otherwise noted.

These cell lines allow the production and propagation of adenovirus-based gene delivery vectors that have deletions in preselected gene regions and that are obtained by cellular complementation of adenoviral genes. To provide the desired complementation of such deleted adenoviral genomes in order to generate a viral vector, plasmid vectors that contain preselected functional units have been designed. Such units include but are not limited to E1 early region, E4 and the viral fiber gene. The preparation of plasmids providing such complementation, thereby being "complementary plasmids or constructs," that are stably inserted into host cell chromosomes are described below.

A. Preparation of an E4-Expressing Plasmid for Complementation of E4-Gene-Deleted Adenoviruses

The viral E4 regulatory region contains a single transcription unit that is alternately spliced to produce several different mRNA products. The E4-expressing plasmid prepared as described herein and used to transfect the 293 cell line contains the entire E4 transcription unit. A DNA fragment extending from 175 nucleotides upstream of the E4 transcription start site including the natural E4 promoter to 153 nucleotides downstream of the E4 polyadenylation signal including the natural E4 terminator signal, corresponding to nucleotides 32667-35780 of the adenovirus type 5 (hereinafter referred to as Ad5) genome as described in Chroboczek *et al.* (*Virology*, 186:280-285 (1992), GenBank Accession Number M73260), was amplified from Ad5 genomic DNA, obtained from the ATCC, via the polymerase chain reaction (PCR). Sequences of the primers used were 5'CGGTACACAGAATTCAGGAGACACAACCTCC3' (forward or 5' primer referred to as E4L) (SEQ ID NO: 1) and 5'GCCTGGATCCGGGAAGTTACGTAACTGGGAAAAC3' (SEQ ID NO: 2) (backward or 3' primer referred to as E4R). To facilitate cloning of the PCR fragment, these oligonucleotides were designed to create new sites for the restriction enzymes EcoRI and BamHI, respectively, as indicated with underlined nucleotides. DNA was amplified via PCR using 30 cycles of 92 C for 1 minute,

-52-

50 C for 1 minute, and 72 C for 3 minutes resulting in amplified full-length E4 gene products.

The amplified DNA E4 products were then digested with EcoRI and BamHI for cloning into the compatible sites of pBluescript/SK+ by standard techniques to create the plasmid pBS/E4. A 2603 base pair (bp) cassette including the herpes simplex virus thymidine kinase promoter, the hygromycin resistance gene, and the thymidine kinase polyadenylation signal was excised from the plasmid pMEP4 (Invitrogen, San Diego, CA) by digestion with FspI followed by addition of BamHI linkers (5'CGCGGATCCGCG3') (SEQ ID NO: 3) for subsequent digestion with BamHI to isolate the hygromycin-containing fragment.

The isolated BamHI-modified fragment was then cloned into the BamHI site of pBS/E4 containing the E4 region to create the plasmid pE4/Hygro containing 8710 bp. The pE4/Hygro plasmid has been deposited with the ATCC under accession number 97739. The complete nucleotide sequence of pE4/Hygro is set forth in SEQ ID NO: 4. Position number 1 of the linearized vector corresponds to approximately the middle portion of the pBS/SK+ backbone. The 5' and 3' ends of the E4 gene are located at respective nucleotide positions 3820 and 707 of SEQ ID NO: 4 while the 5' and 3' ends of the hygromycin insert are located at respective nucleotide positions 3830 and 6470. In the clone that was selected for use, the E4 and hygromycin resistance genes were divergently transcribed.

B. Preparation of a Fiber-Expressing Plasmid for Complementation of Fiber-Gene-Deleted Adenoviruses

To prepare a fiber-encoding construct, primers were designed to amplify the fiber coding region from Ad5 genomic DNA with the addition of unique BamHI and NotI sites at the 5' and 3' ends of the fragment, respectively. The Ad5 nucleotide sequence is available with the GenBank Accession Number M18369. The 5' and 3' primers had the respective nucleotide sequences of 5'ATGGGATCCAAGATGAAGCGCGCAAGACCG3' (SEQ ID NO: 5) and 5'CATAACGCGGCCGCTTCTTTATTCTTGGGC3' (SEQ ID NO: 6), where the inserted BamHI and NotI sites are indicated by underlining. The 5' primer also

-53-

contained a nucleotide substitution 3 nucleotides 5' of the second ATG codon (C to A) that is the initiation site. The nucleotide substitution was included so as to improve the consensus for initiation of fiber protein translation.

5 The amplified DNA fragment was inserted into the BamHI and NotI sites of pcDNA3 (Invitrogen) to create the plasmid designated pCDNA3/Fiber having 7148 bp. The parent plasmid contained the CMV promoter, the bovine growth hormone (BHG) terminator and the gene for conferring neomycin resistance. The viral sequence included in this construct corresponds to nucleotides 31040-32791 of the Ad5 genome.

10 The complete nucleotide sequence of pCDNA3/Fiber is listed in SEQ ID NO: 7 where the nucleotide position 1 corresponds to approximately the middle of the pcDNA3 vector sequence. The 5' and 3' ends of the fiber gene are located at respective nucleotide positions 916 with ATG and 2661 with TAA.

To enhance expression of fiber protein by the constitutive CMV promoter
15 provided by the pcDNA vector, a BglII fragment containing the tripartite leader (TPL) of adenovirus type 5 was excised from pRD112a (Sheay *et al.*, *BioTechniques*, 15:856-862 (1993) and inserted into the BamHI site of pCDNA3/Fiber to create the plasmid pCLF having 7469 bp. The adenovirus tripartite leader sequence, present at the 5' end of all major late adenoviral
20 mRNAs as described by Logan *et al.*, *Proc. Natl. Acad. Sci., USA*, 81:3655-3659 (1984) and Berkner, *BioTechniques*, 6:616-629 (1988), also referred to as a "partial TPL" since it contains a partial exon 1, shows correspondence with the Ad5 leader sequence having three spatially separated exons corresponding to nucleotide positions 6081-6089 (the 3' end of the first leader segment), 7111-
25 7182 (the entire second leader segment), and 9644-9845 (the third leader segment and sequence downstream of that segment). The corresponding cDNA sequence of the partial tripartite leader sequence present in pCLF is included in SEQ ID NO: 8 bordered by BamHI/BglII 5' and 3' sites at respective nucleotide positions 907-912 to 1228-1233. The nucleotide sequence of an isolated partial
30 TPL is also listed separately as SEQ ID No. 22 with the noted 5' and 3' restriction sites and with the following nucleotide regions identified: 1-6 nt BglII site; 1-18 nt polylinker; 19-27 nt last 9 nt of the first leader segment (exon 1);

-54-

28-99 nt second leader segment (exon 2); 100-187 nt third leader segment (exon 3); 188-301 nt contains the nt sequence immediately following the third leader in the genome with an unknown function; and 322-327 nt BglII site.

The pCLF plasmid has been deposited with the ATCC as described in Example 4. The complete nucleotide sequence of pCLF is listed in SEQ ID NO: 8 where the nucleotide position 1 corresponds to approximately the middle of the pcDNA3 parent vector sequence. The 5' and 3' ends of the Ad5 fiber gene are located at respective nucleotide positions 1237-1239 with ATG and 2980-2982 with TAA.

10 C. **Generation of an Adenovirus Packaging Cell Line Carrying Plasmids Encoding Functional E4 and Fiber Proteins**

The 293 cell line was selected for preparing the first adenovirus packaging line as it already contains the E1 gene as prepared by Graham *et al.*, *J. Gen. Virol.*, 36:59-74 (1977) and as further characterized by Spector, *Virol.*, 15 130:533-538 (1983). Before electroporation, 293 cells were grown in RPMI medium + 10% fetal calf serum. Four x 10⁶ cells were electroporated with 20 µg each of pE4/Hygro DNA and pCLF DNA using a BioRad GenePulser and settings of 300 V, 25 µF. DNA for electroporation was prepared using the Qiagen system according to the manufacturer's instructions (Bio-Rad, Richmond, 20 CA).

Following electroporation, cells were split into fresh complete DMEM + 10% fetal calf serum containing 200 µg/ml Hygromycin B (Sigma, St. Louis, MO).

From expanded colonies, genomic DNA was isolated using the 25 "MICROTURBOGEN" system (Invitrogen) according to manufacturer's instructions. The presence of integrated E4 DNA was assessed by PCR using the primer pair E4R and ORF6L (5'TGCTTAAGCGGCCGCGAAGGAGA AGTCC3') (SEQ ID NO: 9), the latter of which is a 5' forward primer near adenovirus 5 open reading frame 6.

30 One clone, designated 211, was selected exhibiting altered growth properties relative to that seen in parent cell line 293. The 211 clone contained the product, indicating the presence of inserted DNA corresponding to most, if

-55-

not all, of the E4 fragment contained in the pE4/Hygro plasmid. The 211 cell line has been deposited with the ATCC as described in Example 4. This line was further evaluated by amplification using the primer pair E4L/E4R described above, and a product corresponding to the full-length E4 insert was detected.

- 5 Genomic Southern blotting was performed on DNA restricted with EcoRI and BamHI. The E4 fragment was then detected at approximately one copy/genome compared to standards with the EcoRI/BamHI E4 fragment as cloned into pBS/E4 for use as a labeled probe with the Genius system according to manufacturer's instructions (Boehringer Mannheim, Indianapolis, IN). In DNA from the 211 cell
- 10 line, the labeled internal fragment pE4/Hygro hybridized with the isolated E4 sequences. In addition, the probe hybridized to a larger fragment which may be the result of a second insertion event.

- Although the 211 cell line was not selected by neomycin resistance, thus indicating the absence of fiber gene, to confirm the lack of fiber gene, the 211
- 15 cell line was analyzed for expression of fiber protein by indirect immunofluorescence with an anti-fiber polyclonal antibody and a FITC-labeled anti-rabbit IgG (KPL) as secondary. No immunoreactivity was detected. Therefore, to generate 211 clones containing recombinant fiber genes, the 211 clone was expanded by growing in RPMI medium and subjected to additional
- 20 electroporation with the fiber-encoding pCLF plasmid as described above.

- Following electroporation, cells were plated in DMEM + 10% fetal calf serum and colonies were selected with 200 μ g/ml G418 (Gibco, Gaithersburg, MD). Positive cell lines remained hygromycin resistant. These candidate
- 25 sublines of 211 were then screened for fiber protein expression by indirect immunofluorescence as described above. The three sublines screened, 211A, 211B and 211R, along with a number of other sublines, all exhibited nuclear staining qualitatively comparable to the positive control of 293 cells infected with AdRSV β gal (1 pfu/cell) and stained 24 hours post-infection.

- Lines positive for nuclear staining in this assay were then subjected to
- 30 Western blot analysis under denaturing conditions using the same antibody. Several lines in which the antibody detected a protein of the predicted molecular weight (62 kd for the Ad5 fiber protein) were selected for further study including

-56-

211A, 211B and 211R. The 211A cell line has been deposited with ATCC as described in Example 4.

Immunoprecipitation analysis using soluble nuclear extracts from these three cell lines and a semipreparative electrophoresis system demonstrated that the fiber protein expressed is in the functional trimeric form characteristic of the native fiber protein. The predicted molecular weight of a trimerized fiber is 186 kd. Under denaturing conditions, the trimeric form was destroyed resulting in detectable fiber monomers. Those clones containing endogenous E1, newly expressed recombinant E4 and fiber proteins were selected for use in complementing adenovirus gene delivery vectors having the corresponding adenoviral genes deleted as described in Example 2.

D. Preparation of an E1-Expressing Plasmid for Complementation of E1-Gene-Deleted Adenoviruses

In order to prepare adenoviral packaging cell lines other than those based on the E1-gene containing 293 cell line as described in Example 1C above, plasmid vectors containing E1 alone or in various combinations with E4 and fiber genes are constructed as described below.

The region of the adenovirus genome containing the E1a and E1b gene is amplified from viral genomic DNA by PCR as previously described. The primers used are E1L, the 5' or forward primer, and E1R, the 3' or backward primer, having the respective nucleotide sequences 5'CCG AGCTAGC GACTGAAAATGAG3' (SEQ ID NO: 10) and 5'CCTCTCGAG AGACAGC AAGACAC3' (SEQ ID NO: 11). The E1L and E1R primers include the respective restriction sites NheI and XhoI as indicated by the underlines. The sites are used to clone the amplified E1 gene fragment into the NheI/XhoI sites in pMAM commercially available from Clontech (Palo Alto, CA) to form the plasmid pDEX/E1 having 11152 bp.

The complete nucleotide sequence of pDEX/E1 is listed in SEQ ID NO: 12 where the nucleotide position 1 corresponds to approximately 1454 nucleotides from the 3' end of the pMAM backbone vector sequence. The pDEX/E1 plasmid includes nucleotides 552 to 4090 of the adenovirus genome positioned downstream (beginning at nucleotide position 1460 and ending at 4998 in the

-57-

pDEX/E1 plasmid) of the glucocorticoid-inducible mouse mammary tumor virus (MMTV) promoter of pMAM. The pMAM vector contains the *E. coli gpt* gene that allows stable transfectants to be isolated using hypoxanthine/aminopterin/thymidine (HAT) selection. The pMAM backbone occupies nucleotide positions 1-1454 and 5005-11152 of SEQ ID NO: 12.

E. Generation of an Adenovirus Packaging Cell Line Carrying Plasmids Encoding Functional E1, and Fiber Proteins

To create separate adenovirus packaging cell lines equivalent to that of the 211 sublines, 211A, 211B and 211R, as described in Example 1C, alternative cell lines lacking adenoviral genomes are selected for transfection with the plasmid constructs as described below. Acceptable host cells include A549, Hela, Vero and the like cell lines as described in Example 1. The selected cell line is transfected with the separate plasmids, pDEX/E1 and pCLF, respectively for expressing E1, and fiber complementary proteins. Following transfection procedures as previously described, clones containing stable insertions of the two plasmids are isolated by selection with neomycin and HAT. Integration of full-length copy of the E1 gene is assessed by PCR amplification from genomic DNA using the primer set E1L/E1R, as described above. Functional insertion of the fiber gene is assayed by staining with the anti-fiber antibody as previously described.

The resultant stably integrated cell line is then used as a packaging cell system to complement adenoviral gene delivery vectors having the corresponding adenoviral gene deletions as described in Example 2.

F. Preparation of a Plasmid Containing Two or More Adenoviral Genes for Complementing Gene-Deleted Adenoviruses

The methods described in the preceding Examples rely on the use of two plasmids, pE4/Hygro and pCLF, or, pCLF and pDEX/E1 for generating adenoviral cell packaging systems. In alternative embodiments, complementing plasmids containing two or more adenoviral genes for expressing of encoded proteins in various combinations are also prepared as described below. The resultant plasmids are then used in various cell systems with delivery plasmids having the corresponding adenoviral gene deletions. The selection of packaging cell, content of the delivery plasmids and content of the complementing plasmids for

-58-

use in generating recombinant adenovirus viral vectors thus depends on whether other adenoviral genes are deleted along with the adenoviral fiber gene, and, if so, which ones.

5 1. **Preparation of a Complementing Plasmid Containing Fiber and E1 Adenoviral Genes**

A DNA fragment containing sequences for the CMV promoter, adenovirus tripartite leader, fiber gene and bovine growth hormone terminator is amplified from pCLF prepared in Example 1B using the forward primer 5'GACGGATCGGGAGATCTCC3' (SEQ ID NO: 13), that anneals to the
10 nucleotides 1-19 of the pCDNA3 vector backbone in pCLF, and the backward primer 5'CCGCCTCAGAAGCCATAGAGCC3' (SEQ ID NO: 14) that anneals to nucleotides 1278-1257 of the pCDNA3 vector backbone. The fragment is amplified as previously described and then cloned into the pDEX/E1 plasmid, prepared in Example 1D. For cloning in the DNA fragment, the pDEX/E1 vector
15 is first digested with NdeI, that cuts at a unique site in the pMAM vector backbone in pDEX/E1, then the ends are repaired by treatment with bacteriophage T4 polymerase and dNTPs.

The resulting plasmid containing E1 and fiber genes, designated pE1/Fiber, provides dexamethasone-inducible E1 function as described for
20 DEX/E1 and expression of Ad5 fiber protein as described above.

The complete nucleotide sequence of pE1/Fiber is listed in SEQ ID NO: 15 where the nucleotide position 1 corresponds to approximately 1459 nucleotides from the 3' end of the parent vector pMAM sequence. The 5' and 3' ends of the Ad5 E1 gene are located at respective nucleotide positions 1460 and 4998
25 followed by pMAM backbone and then separated from the Ad5 fiber from pCLF by the filled-in blunt ended NdeI site. The 5' and 3' ends of the pCLF fiber gene fragment are located at respective nucleotide positions 10922-14223 containing elements as previously described for pCLF.

The resultant pE1/Fiber plasmid is then used to complement one or more
30 delivery plasmids expressing E1 and fiber.

The pE1/Fiber construct is then used to transfect a selected host cell as described in Example 1E to generate stable chromosomal insertions performed as

-59-

previously described followed by selection on HAT medium. The stable cells are then used as packaging cells as described in Example 2.

2. Preparation of a Complementing Plasmid Containing E4 and Fiber Adenoviral Genes

5 Plasmid pCLF prepared as described in Example 1B is partially digested with BglII to cut only at the site in the pCDNA3 backbone. The pE4/Hygro plasmid prepared in Example 1A is digested with BamHI to produce a fragment containing E4. The E4 fragment is then inserted into the BamHI site of pCLF to form plasmid pE4/Fiber. The resultant plasmid provides expression of the fiber
10 gene as described for pCLF and E4 function as described for pE4/Hygro.

A schematic plasmid map of pE4/Fiber, having 10610 bp. The complete nucleotide sequence of pE4/Fiber is listed in SEQ ID NO: 16 where the nucleotide position 1 corresponds to approximately 14 bp from the 3' end of the parent vector pCDNA3 backbone sequence. The 5' and 3' ends of the Ad5 E4
15 gene are located at respective nucleotide positions 21 and 3149 followed by fused BglII/BamHI sites and pCDNA3 backbone including the CMV promoter again followed by BglII/BamHI sites. The adenovirus leader sequence begins at nucleotide position 4051 and extends to 4366 followed by fused BamHI/BglII sites and the 5' and 3' ends of the fiber gene located at respective nucleotide
20 positions 4372 and 6124.

Stable chromosomal insertions of pE4/Fiber in host cells are obtained as described above.

EXAMPLE 2

25 Preparation of Adenoviral Gene Delivery Vectors Using Adenoviral Packaging Cell Lines

Adenoviral delivery vectors are prepared to separately lack the combinations of E1/fiber and E4/fiber. Such vectors are more replication-defective than those previously in use due to the absence of multiple viral genes. A preferred adenoviral delivery vector is replication competent but only via a
30 non-fiber means is one that only lacks the fiber gene but contains the remaining functional adenoviral regulatory and structural genes. Furthermore, these adenovirus delivery vectors have a higher capacity for insertion of foreign DNA.

-60-

A. Preparation of Adenoviral Gene Delivery Vectors Having Specific Gene Deletions and Methods of Use

- To construct the E1/ fiber deleted viral vector containing the LacZ reporter gene construct, two new plasmids were constructed. The plasmid pΔ E1Bβgal was constructed as follows. A DNA fragment containing the SV40 regulatory sequences and *E. coli* β-galactosidase gene was isolated from pSVβgal (Promega) by digesting with VspI, filling the overhanging ends by treatment with Klenow fragment of DNA polymerase I in the presence of dNTP's and digesting with Bam HI. The resulting fragment was cloned into the EcoRV and BamHI sites in the polylinker of pΔ E1sp1B (Microbix Biosystems, Hamilton, Ontario) to form pΔ E1B βgal that therefore contained the left end of the adenovirus genome with the Ela region replaced by the LacZ cassette (nucleotides 6690 to 4151) of pSVβgal. Plasmid DNA may be prepared by the alkaline lysis method as described by Birnboim and Doly, *Nuc. Acids Res.*, 7:1513-1523 (1978) or by the QuiaGen method according to the manufacturer's instruction, from transformed cells used to expand the plasmid DNA was then purified by CsCl-ethidium bromide density gradient centrifugation. Alternatively, plasmid DNAs may be purified from *E. coli* by standard methods known in the art (e.g. see Sambrook *et al.*)
- The second plasmid (pDV44), prepared as described herein, is derived from pBHG10, a vector prepared as described by Bett *et al.*, *Proc. Natl. Acad. Sci., USA*, 91:8802-8806 (1994) (see, also International PCT application No. WO 95/00655) using methods well known to one of skill in the art. This vector is also commercially available from Microbix and contains an Ad5 genome with the packaging signals at the left end deleted and the E3 region (nucleotides 28133:30818) replaced by a linker with a unique site for the restriction enzyme PacI. An 11.9 kb BamHI fragment, which contains the right end of the adenovirus genome, is isolated from pBHG10 and cloned into the BamHI site of pBS/SK(+) to create plasmid p11.3 having approximately 14,658 bp. The p11.3 plasmid was then digested with PacI and Sall to remove the fiber, E4, and inverted terminal repeat (ITR) sequences.

-61-

This fragment was replaced with a 3,4 kb fragment containing the ITR segments and the E4 gene which was generated by PCR amplification from pBHG10 using the following oligonucleotide sequences:

5' TGTACACCG GATCCGGCGCACACC3' SEQ ID NO: 17; and

5 5' CACAACGAGCTC AATTAATTAATTGCCACATCCTC3' SEQ ID NO: 18.

These primers incorporated sites for *PacI* and *BamHI*. Cloning this fragment into the *PacI* and blunt ended *Sall* sites of the p11.3 backbone resulted in a substitution of the fused ITRs, E4 region and fiber gene present in pBHG10, by the ITRs and E4 region alone. The resulting p11.3 plasmid containing the ITR and E4 regions, designated plasmid pDV43a, was then digested with *BamHI*. This *BamHI* fragment was then used to replace a *BamHI* fragment in pBHG10 thereby creating pDV44 in a pBHG10 backbone.

In an alternative approach to preparing pDV44 with an additional subcloning step to facilitate the incorporation of restriction cloning sites, the following cloning procedure was performed. pDV44 as above was constructed by removing the fiber gene and some of the residual E3 sequences from pBHG10 (Microbix Biosystems). As above, to simplify manipulations, the 11.9 kb *BamHI* fragment including the rightmost part of the Ad5 genome was removed from pBHG10 and inserted into pBS/SK. The resulting plasmid was termed p11.3. The 3.4 kb DNA fragment corresponding to the E4 region and both ITRs of adenovirus type 5 was amplified as described above from pBHG10 using the oligonucleotides listed above and subcloned into the vector pCR2.1 (Invitrogen) to create pDV42. This step is the additional cloning step to facilitate the incorporation of a *Sall* restriction site. pDV42 was then digested with *PacI*, which cuts at a unique site (bold type) in one of the PCR primers, and with *Sall*, which cuts at a unique site in the pCR2.1 polylinker. This fragment was used to replace the corresponding *PacI/XhoI* fragment of p11.3 (the pBS polylinker adjacent to the Ad DNA fragment contains a unique *XhoI* site), creating pDV43.

30 A plasmid designated pDV44 was constructed by replacing the 11.9 kb *BamHI* fragment of pBHG10 by the analogous *BamHI* fragment of pDV43. As generated in the first procedure, pDV44 therefore differs from pBHG10 by the

-62-

deletion of Ad5 nucleotides 30819:32743 (residual E3 sequences and all but the 3'-most 41 nucleotides of the fiber open reading frame).

Thus, to summarize, the cloning procedures described above result in the production of a fiber-deleted Ad5 genomic plasmid (pDV44) that was constructed by removing the fiber gene and some of the residual E3 sequences from pBHG10. pDV44 contains a wild-type E4 region, but only the last 41 nucleotides of the fiber ORF (this sequence was retained to avoid affecting expression of the adjacent E4 transcription unit). Plasmids pBHG10 and pDV44 contain unpackageable Ad5 genomes, and must be rescued by cotransfection and subsequent homologous recombination with DNA carrying functional packaging signals. In order to generate vectors marked with a reporter gene, either pDV44 or pBHG10 was cotransfected with p Δ E1B β gal, which contains the left end of the Ad5 genome with an SV40-driven β -galactosidase reporter gene inserted in place of the E1 region.

In general, and as described below, the method for virus production by recombination of plasmids followed by complementation in cell culture involves the isolation of recombinant viruses by cotransfection of any one of the adenovirus packaging cell systems prepared in Example 1, namely 211A, 211B, 211R, A549, Vero cells, and the like, with plasmids carrying sequences corresponding to viral gene delivery vectors.

A selected cell line is plated in dishes and cotransfected with pDV44 and p Δ E1B β gal using the calcium phosphate method as described by Bett *et al.*, *Proc. Natl. Acad. Sci., USA*, 91:8802-8806 (1994). Recombination between the overlapping adenovirus sequences in the two plasmids leads to the creation of a full-length viral chromosome where pDV44 and p Δ E1B β gal recombine to form a recombinant adenovirus vector having multiple deletions. The deletion of E1 and of the fiber gene from the viral chromosome is compensated for by the sequences integrated into the packaging cell genome, and infectious virus particles are produced. The plaques thus generated are isolated and stocks of the recombinant virus are produced by standard methods.

Because of the fiber deletion, a pDV44-derived virus is replication-defective, cells in which it is grown must complement this defect.

-63-

The 211B cell line (a derivative of 293 cells which expresses the wild-type (wt) AD5 fiber and is equivalent to 211A on deposit with ATCC as described in Example 4) was used for rescue and propagation of the virus described here. pDV44 and pΔE1βgal were cotransfected into 211B cells, and the monolayers
5 were observed for evidence of cytopathic effect (CPE). Briefly, for virus construction, cells were transfected with the indicated plasmids using the Gibco Calcium Phosphate Transfection system according to the manufacturer's instructions and observed daily for evidence of CPE.

One of a total of 58 transfected dishes showed evidence of spreading cell
10 death at day 15. A crude freeze-thaw lysate was prepared from these cells and the resulting virus (termed Ad5.βgal.ΔF) was plaque purified twice and then expanded. To prepare purified viral preparations, cells were infected with the indicated Ad and observed for completion of CPE. Briefly, at day zero, 211B cells were plated in DMEM plus 10% fetal calf serum at approximately 1×10^7
15 cells/150 cm² flask or equivalent density. At day one, the medium was replaced with one half the original volume of fresh DMEM containing the indicated Ad, in this case Ad5.βgal.ΔF, at approximately 100 particles/cell. At day two, an equal volume of medium was added to each flask and the cells were observed for CPE. Two to five days after infection, cells were collected and virus isolated by lysis
20 via four rapid freeze-thaw cycles. Virus was then purified by centrifugation on preformed 15-40% CsCl gradients (111,000 x g for three hours at 4°C). The bands were harvested, dialyzed into storage buffer (10 mM Tris-pH 8.1, 0.9% NaCl, and 10% glycerol), aliquoted and stored at -70°C. Purified Ad5.βgal.ΔF virus particles containing human adenovirus Ad5.βgal.ΔF genome (described
25 further below) have been deposited with the ATCC on January 15, 1999 as further described in Example 4.

For viral titering, as necessary in the below Examples, Ad preparations were titrated by plaque assay on 211B cells. Cells were plated on
polylysine-coated 6 well plates at 1.5×10^6 cells/well. Duplicate dilutions of
30 virus stock were added to the plates in 1 ml/well of complete DMEM. After a five hour incubation at 37°C, virus was removed and the wells overlaid with 2

-64-

ml of 0.6% low-melting agarose in Medium 199 (Gibco). An additional 1 ml of overlay was added at five day intervals.

As a control, the first-generation virus Ad5. β gal.wt, which is identical to Ad5. β gal. Δ F except for the fiber deletion, was constructed by cotransfection of pBHG10 and p Δ E1 β gal. In contrast to the low efficiency of recovery of the fiberless genome (1/58 dishes), all of 9 dishes cotransfected with p Δ E1 β gal and pBHG10 produced virus.

In another embodiment, a delivery plasmid is prepared that does not require the above-described recombination events to prepare a viral vector having a fiber gene deletion. In one embodiment, a single delivery plasmid containing all the adenoviral genome necessary for packaging but lacking the fiber gene is prepared from plasmid pFG140 containing full-length Ad5 that is commercially available from Microbix. The resultant delivery plasmid referred to as pFG140-f is then used with pCLF stably integrated cells as described above to prepare a viral vector lacking fiber. For genetic therapy, the fiber gene can be replaced with a therapeutic gene of interest for preparing a therapeutic delivery adenoviral vector. Methods for producing a fiberless vector with a complete TPL are described in Example 3.

Vectors for the delivery of any desired gene and preferably a therapeutic gene are prepared by cloning the gene of interest into the multiple cloning sites in the polylinker of commercially available p Δ E1sp1B (Microbix Biosystems), in an analogous manner as performed for preparing pE1 β gal as described above. The same cotransfection and recombination procedure is then followed as described herein to obtain viral gene delivery vectors as further discussed in later Examples.

1. Characterization of the Ad5. β gal. Δ F Genome

To confirm that the vector genomes had the proper structures and that the fiber gene was absent from the Ad5. β gal. Δ F chromosome, the DNA isolated from viral particles was analyzed. Briefly, purified viral DNA was obtained by adding 10 μ l of 10 mg/ml proteinase K, 40 μ l of 0.5 M EDTA and 50 μ l of 10% SDS to 800 μ l of adenovirus-containing culture supernatant. The suspension

-65-

was then incubated at 55°C for 60 minutes. The solution was then extracted once with 400 μ l of a 24:1 mixture of chloroform:isoamyl alcohol. The aqueous phase was then removed and precipitated with sodium acetate/ethanol. The pellet was washed once with 70% ethanol and lightly dried. The pellet was then suspended in 40 μ l of 10 mM Tris-HCl, pH 8.0, 1 mM EDTA. Genomic DNA from Ad5. β gal.wt and Ad5. β gal. Δ F produced the expected restriction patterns following digestion with either EcoRI or with NdeI. Southern blotting, performed with standard methods, with labeled fiber DNA as a probe demonstrated the presence of fiber sequence in Ad5. β gal.wt but not in Ad5. β gal. Δ F DNA. As a positive control, the blot was stripped and reprobed with labeled E4 sequence. Fiber and E4 sequences were detected by using labeled inserts from pCLF and pE4/Hygro, respectively. E4 signal was readily detectable in both genomes at equal intensities. The complete nucleotide sequence of Ad5. β gal. Δ F is presented in SEQ ID NO: 23 and is contained in the virus particle on deposit with ATCC.

2. Characterization of the Fiberless Adenovirus Ad5. β gal. Δ F

To verify that Ad5. β gal. Δ F was fiber-defective, 293 cells (which are permissive for growth of E1-deleted Ad vectors but do not express fiber) were infected with Ad5. β gal. Δ F or with Ad5. β gal.wt. Twenty-four hours post infection, the cells were stained with polyclonal antibodies directed either against fiber or against the penton base protein. Cells infected with either virus were stained by the anti-penton base antibody, while only cells infected with the Ad5. β gal.wt control virus reacted with the anti-fiber antibody. This confirms that the fiber-deleted Ad mutant does not direct the synthesis of fiber protein.

3. Growth of the Fiber-Deleted Ad5. β gal. Δ F Vector in Complementing Cells

Ad5. β gal. Δ F was found to readily be propagated in 211B cells. As assayed by protein concentration, CsCl-purified stocks of either Ad5. β gal. Δ F or Ad5. β gal.wt contained similar numbers of viral particles. The particles appeared to band normally on CsCl gradients. Infectivity of the Ad5. β gal. Δ F particles was lower than the Ad5. β gal.wt control, as indicated by an increased particle/PFU ratio. Ad5. β gal. Δ F was also found to plaque more slowly than the control

-66-

virus. When plated on 211B cells, Ad5. β gal.wt plaques appeared within 5-7 days, while plaques of Ad5. β gal. Δ F continued to appear until as much as 15-18 days post infection. Despite their slower formation, the morphology of Ad5. β gal. Δ F plaques was essentially normal.

5 4. Production of Fiberless Ad5. β gal. Δ F Particles

As Ad5. β gal. Δ F represents a true fiber null mutation and its stocks are free of helper virus, the fiber mutant phenotype was readily investigated. A single round of growth in cells (such as 293) which do not produce fiber generating a homogeneous preparation of fiberless Ad allowed for the
10 determination of whether such particles would be stable and/or infectious. Either Ad5. β gal.wt or Ad5. β gal. Δ F was grown in 293 or 211B cells, and the resulting particles purified on CsCl gradients as previously described. Ad5. β gal. Δ F particles were readily produced in 293 cells at approximately the same level as the control virus and behaved similarly on the gradients, indicating
15 that there was not a gross defect in morphogenesis of fiberless capsids.

Particles of either virus contained similar amounts of penton base regardless of the cell type in which they were grown. This demonstrated that fiber is not required for assembly of the penton base complex into virions. The Ad5. β gal. Δ F particles produced in 293 cells did not contain fiber protein.
20 211B-grown Ad5. β gal. Δ F also contained less fiber than the Ad5. β gal.wt control virus. The infectivities of the different viral preparations on epithelial cells correlated with the amount of fiber protein present. The fiberless Ad particles were several thousand-fold less infectious than the first-generation vector control on a per-particle basis, while infectivity of 211B-grown Ad5. β gal. Δ F was only
25 50-100 fold less than that of Ad5. β gal.wt. These studies confirmed fiber's crucial role in infection of epithelial cells via CAR binding.

5. Composition and Structure of the Fiberless Ad5. β gal. Δ F Particles

The proteins contained in particles of 293-grown Ad5. β gal. Δ F were
30 compared to those in Ad5. β gal.wt, to determine whether proteolysis or particle assembly was defective in this fiber null mutant. The overall pattern of proteins in the fiberless particles was observed to be quite similar to that of a

-67-

first-generation vector, with the exception of reduced intensity of the composite band resulting from proteins IIIa and IV (fiber). The fiberless particles also had a reduced level of protein VII. Although substantial amounts of uncleaved precursors to proteins VI, VII, and VIII were not seen, it is possible that the

5 low-molecular weight bands migrating ahead of protein VII represent either aberrantly cleaved viral proteins or their breakdown products.

Cryo-electron microscopy was used to more closely examine the structure of the 293 grown Ad5.βgal.ΔF and of Ad5βgal.wt. The fiber, having an extended stalk with a knob at the end, was faintly visible in favorable

10 orientations of wild-type Ad5 particles, but not in images of the fiberless particles. Filamentous material likely corresponding to free viral DNA was seen in micrographs of fiberless particles. This material was also present in micrographs of the first-generation control virus, albeit at much lower levels.

Three-dimensional image reconstructions of fiberless and wild-type

15 particles at ~20 Å resolution showed similar sizes and overall features, with the exception that fiberless particles lacked density corresponding to the fiber protein. The densities corresponding to other capsid proteins, including penton base and proteins IIIa, VI, and IX, were comparable in the two structures. This confirms that absence of fiber does not prevent assembly of these components

20 into virions. The fiber was truncated in the wild-type structure as only the lower portion of its flexible shaft follows icosahedral symmetry. The RGD protrusions on the fiberless penton base were angled slightly inward relative to those of the wild-type structure. Another difference between the two penton base proteins was that there is a ~30 Å diameter depression in the fiberless penton base

25 around the five-fold axis where the fiber would normally sit. The Ad5 reconstructions confirm that capsid assembly, including addition of penton base to the vertices, is able to proceed in the complete absence of fiber.

6. Integrin-Dependent Infectivity of Fiberless Ad5.βgal.ΔF Particles

30 While attachment via the viral fiber protein is a critical step in the infection of epithelial cells, an alternative pathway for infection of certain hematopoietic cells has been described. In this case, penton base mediates

-68-

binding to the cells (via $\beta 2$ integrins) and internalization (through interaction with α integrins). Particles lacking fiber might therefore be expected to be competent for infection of these cells, even though on a per-particle basis they are several thousand-fold less infectious than normal Ad vectors on epithelial cells.

To investigate this, THP-1 monocytic cells were infected with Ad5. β gal.wt or with Ad5. β gal. Δ F grown in the absence of fiber. Infection of THP-1 cells was assayed by infecting 2×10^5 cells at the indicated m.o.i. in 0.5 ml of complete RPMI. Forty-eight hours post-infection, the cells were fixed with glutaraldehyde and stained with X-gal, and the percentage of stained cells was determined by light microscopy. The results of the infection assay showed that the fiberless particles were only a few-fold less infectious than first-generation Ad on THP-1 cells. Large differences were seen in plaquing efficiency on epithelial (211B) cells. Infection of THP-1 cells by either Ad5. β gal. Δ F or Ad5. β gal.wt was not blocked by an excess of soluble recombinant fiber protein, but could be inhibited by the addition of recombinant penton base). These results indicate that the fiberless Ad particles use a fiber-independent pathway to infect these cells. Furthermore, the lack of fiber protein did not prevent Ad5. β gal. Δ F from internalizing into the cells and delivering its genome to the nucleus, demonstrating that fiberless particles are properly assembled and are capable of uncoating.

The foregoing results with the recombinant viruses thus produced indicates that they can be used as gene delivery tools in cultured cells and *in vivo* as described more fully in the Examples. For example, for studies of the effectiveness and relative immunogenicity of multiply-deleted vectors, virus particles are produced by growth in the packaging lines described in Example 1 and are purified by CsCl gradient centrifugation. Following titering, virus particles are administered to mice via systemic or local injection or by aerosol delivery to lung. The LacZ reporter gene allows the number and type of cells which are successfully transduced to be evaluated. The duration of transgene expression is evaluated in order to determine the long-term effectiveness of treatment with multiply-deleted recombinant adenoviruses relative to the

-70-

M18369 and M12411. Oligonucleotide primers are designed to amplify the entire coding sequence of the full-length fiber genes, starting from the start codon, ATG, and ending with the termination codon TAA. For cloning purposes, the 5' and 3' primers contain the respective restriction sites BamHI and NotI for
5 cloning into pcDNA plasmid as described in Example 1A. PCR is performed as described above.

The resulting products are then used to construct chimeric fiber constructs by PCR gene overlap extension (Horton *et al.* (1990) *BioTechniques*, 8:525-535). The Ad5 fiber tail and shaft regions (5TS; the nucleotide region
10 encoding amino acid residue positions 1 to 403) are connected to the Ad3 fiber head region (3H; the nucleotide region encoding amino acid residue positions 136 to 319) to form the 5TS3H fiber chimera. Conversely, the Ad3 fiber tail and shaft regions (3TS; the nucleotide region encoding amino acid residues positions
15 encoding the amino acid residue positions 404 to 581) to form the 3TS5H fiber chimera. The fusions are made at the conserved TLWT (SEQ ID NO: 19) sequence at the fiber shaft-head junction.

The resultant chimeric fiber PCR products are then digested with BamHI and NotI for separate directional ligation into a similarly digested pcDNA 3.1.
20 The TPL sequence is then subcloned into the BamHI as described in Example 1A for preparing an expression vector for subsequent transfection into 211 cells as described above or into the alternative packaging cell systems as previously described. The resultant chimeric fiber construct-containing adenoviral packaging cell lines are then used to complement adenoviral delivery vectors as
25 previously described. Other fiber chimeric constructs are obtained with the various adenovirus serotypes using a similar approach.

In an alternative embodiment, the use of modified proteins including with modified epitopes (see, *e.g.*, Michael *et al.* (1995) *Gene Therapy*, 2:660-668 and International PCT application Publication No. WO 95/26412, which describe
30 the construction of a cell-type specific therapeutic viral vector having a new binding specificity incorporated into the virus concurrent with the destruction of the endogenous viral binding specificity). In particular, the authors described the

-71-

production of an adenoviral vector encoding a gastrin releasing peptide (GRP) at the 3' end of the coding sequence of the Ad5 fiber gene. The resulting fiber-GRP fusion protein was expressed and shown to assemble functional fiber trimers that were correctly transported to the nucleus of HeLa cells following
5 synthesis.

Similar constructs are contemplated for use in the complementing adenoviral packaging cell systems for generating new adenoviral gene delivery vectors that are targetable, replication-deficient and less immunogenic. Heterologous ligands contemplated for use herein to redirect fiber specificity
10 range from as few as 10 amino acids in size to large globular structures, some of which necessitate the addition of a spacer region so as to reduce or preclude steric hindrance of the heterologous ligand with the fiber or prevent trimerization of the fiber protein. The ligands are inserted at the end or within the linker region. Preferred ligands include those that target specific cell receptors or
15 those that are used for coupling to other moieties such as biotin and avidin.

A preferred spacer includes a short 12 amino acid peptide linker composed of a series of serines and alanine flanked by a proline residue at each end using routine procedures known to those of skill in the art. The skilled artisan will be with the preparation of linkers to accomplish sufficient protein
20 presentation and to alter the binding specificity of the fiber protein without compromising the cellular events that follow viral internalization. Moreover, within the context of this disclosure, preparation of modified fibers having ligands positioned internally within the fiber protein and at the carboxy terminus as described below are contemplated for use with the methods described herein.

25

The preparation of a fiber having a heterologous binding ligand is prepared essentially as described in the above-cited paper. Briefly, for the ligand of choice, site-directed mutagenesis is used to insert the coding sequence for a linker into the 3' end of the Ad5 fiber construct in pCLF as prepared in Example

30 1.

The 3' or antisense or mutagenic oligonucleotide encodes a preferred linker sequence of ProSerAlaSerAlaSerAlaSerAlaProGlySer (SEQ ID NO: 20)

-72-

followed by a unique restriction site and two stop codons, respectively, to allow the insertion of a coding sequence for a selected heterologous ligand and to ensure proper translation termination. Flanking this linker sequence, the mutagenic oligonucleotide contains sequences that overlap with the vector
5 sequence and allow its incorporation into the construct. Following mutagenesis of the pCLF sequence adding the linker and stop codon sequences, a nucleotide sequence encoding a preselected ligand is obtained, linkers corresponding to the unique restriction site in the modified construct are attached and then the sequence is cloned into linearized corresponding restriction site. The
10 resultant fiber-ligand construct is then used to transfect 211 or the alternative cell packaging systems previously described to produce complementing viral vector packaging systems.

In a further embodiment, intact fiber genes from different Ad serotypes are expressed by 211 cells or an alternative packaging system as previously
15 described. A gene encoding the fiber protein of interest is first cloned to create a plasmid analogous to pCLF, and stable cell lines producing the fiber protein are generated as described above for Ad5 fiber. The adenovirus vector described which lacks the fiber gene is then propagated in the cell line producing the fiber
20 protein relevant for the purpose at hand. As the only fiber gene present is the one in the packaging cells, the adenoviruses produced contain only the fiber protein of interest and therefore have the binding specificity conferred by the complementing protein. Such viral particles are used in studies such as those described above to determine their properties in experimental animal systems.

EXAMPLE 3

25 Tripartite leader sequences (TPLs) that are useful in enhancing the expression of complementing adenoviral proteins, particularly fiber protein, for use in preparing an adenoviral gene delivery vector are provided. The complete Ad5 TPL was constructed by assembling PCR fragments. First, the third TPL exon (exon 3) (nt 9644-9731 of the Ad5 genome) was amplified from Ad5
30 genomic DNA using the synthetic oligonucleotide primers
5'CTCAACAATTGTGGATCCGTA CTCC3' (SEQ ID No. 24) and
5'GTGCTCAGCAGATCTTGCGACTGTG3' (SEQ ID No. 25). The resulting

-73-

product was cloned to the BamHI and BglII sites of pΔE1Sp1a (Microbix Biosystems) using sites in the primers (shown in bold) to create plasmid pDV52. A fragment corresponding to the first TPL exon (exon 1), the natural first intron (intron 1), and the second TPL exon (exon 2) (Ad5 nt 6049-7182) was then
5 amplified using primers 5'GGCGCGTTCGGATCCACTCTCTTCC3' (SEQ ID No. 26) and 5'CTACATGCTAGGCAGATCTCGTTCGGAG3' (SEQ ID No. 27), and cloned into the BamHI site of pDV52 (again using sites in the primers) to create pDV55.

This plasmid contains a 1.2 kb BamHI/BglII fragment containing the first
10 TPL exon, the natural first intron, and the fused second and third TPL exons. The nucleotide sequence of the complete TPL containing the noted 5' and 3' restriction sites is shown in SEQ ID No 28 with the following nucleotide regions identified: 1-6 nt BamHI site; 7-47 nt first leader segment (exon 1); 48-1068 nt natural first intron (intron 1); 1069-1140 nt second leader segment (exon 2);
15 1141-1146 nt fused BamHI and BglII sites; 1147-1234 nt third leader segment (exon 3); and 1235-1240 nt BglII site.

EXAMPLE 4

Deposit of Materials

The following cell lines and plasmids were deposited on September 25,
20 1996, with the American Type Culture Collection, 10801 University Blvd, Manassas, Virginia, USA (ATCC) under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty): Plasmid pE4/Hygro (accession number 97739), Plasmid pCLF (accession number 97737),
25 211 Cell Line (accession number CRL-12193) and 211A Cell Line (accession number CRL-12194)

The following virus, Ad5.βgal.ΔF, was deposited on January 15, 1999, with the ATCC as listed above and provided with accession number VR2636.

Additionally, plasmids pDV60, pDV67, pDV69, pDV80 and pDV90 were
30 deposited at the ATCC on January 5, 2000 and provided with accession numbers PTA-1144, PTA-1145, PTA-1146, PTA-1147 and PTA-1148 respectively.

-74-

EXAMPLE 5**Preparation and Use of Adenoviral Packaging Cell Lines
Containing Plasmids Containing Alternative TPLs**

5 Plasmids containing tripartite leaders (TPLs) have been constructed. The resulting plasmids that contain different selectable markers, such as neomycin and zeocin, were then used to prepare fiber-complementing stable cell lines for use as for preparing adenoviral vectors.

A. pDV60

10 Plasmid pDV60 was constructed by inserting this TPL cassette of SEQ ID No. 28 into the BamHI site upstream of the Ad5 fiber gene in pcDNA3/Fiber, a neomycin selectable plasmid (see, *e.g.*, U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265 on January 14, 2000); see also Von Seggern *et al.* (1998) *J. Gen Virol.*, 79: 1461-1468). The nucleotide sequence of pDV60 is
15 listed in SEQ ID NO: 29. Plasmid pDV60 has been deposited in the ATCC under accession number PTA-1144.

B. pDV61

To construct pDV61, an Asp718/NotI fragment containing the CMV promoter, partial Ad5 TPL, wildtype Ad5 fiber gene, and bovine growth hormone
20 terminator was transferred from pCLF (ATCC accession number 97737; and described in copending U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265 on January 14, 2000);), to a zeocin selectable cloning vector referred to as pcDNA3.1/Zeo (+) (commercially available from Invitrogen and for which the sequence is known).

25 C. pDV67

In an analogous process, pDV67 containing complete TPL was constructed by transferring an Asp 718/XbaI fragment from pDV60 into pcDNA3.1/Zeo(+) backbone. The nucleotide sequence of pDV67 is set forth in

-75-

SEQ ID No. 30. Plasmid pDV67 is available from the ATCC under accession number PTA-1145.

D. pDV69

To prepare pDV69 containing a modified fiber protein, the chimeric
5 Ad3/Ad5 fiber gene was amplified from pGEM5TS3H (Stevenson *et al.* (1995) *J. Virol.*, 69: 2850-2857) using the primers 5'ATGGGAT
CAAGATGAAGCGCGCAAGACCG3' (SEQ ID No. 31) and
5'CACTATAGCGGCCGCATTCTCAGTCATCTT3' (SEQ ID No. 32), and cloned to
the BamHI and NotI sites of pcDNA3.1/Zeo(+) via new BamHI and NotI sites
10 engineered into the primers to create pDV68. Finally, the complete TPL
fragment described above was then added to the unique BamHI site of pDV68
to create pDV69. The nucleotide sequence of pDV69 is listed in SEQ ID No. 33
and has been deposited in the ATCC under accession number PTA-1146.

E. Preparation of Stable Adenovirus Packaging Cell Lines

15 E1-2a S8 cells are derivatives of the A549 lung carcinoma line (ATCC #
CCL 185) with chromosomal insertions of the plasmids pGRE5-2.E1 (also
referred to as GRE5-E1-SV40-Hygro construct and listed in SEQ ID No. 34) and
pMNeoE2a-3.1 (also referred to as MMTV-E2a-SV40-Neo construct and listed in
SEQ ID No. 35), which provide complementation of the adenoviral E1 and E2a
20 functions, respectively. This line and its derivatives were grown in Richter's
modified medium (BioWhittaker) + 10% FCS. E1-2a S8 cells were
electroporated as previously described (Von Seggern *et al.* (1998) *J. Gen Virol.*,
79: 1461-1468) with pDV61, pDV67, or with pDV69, and stable lines were
selected with zeocin (600 µg/ml).

25 The cell line generated with pDV61 is designated 601. The cell line
generated with pDV67 is designated 633 while that generated with pDV69 is
designated 644. Candidate clones were evaluated by immunofluorescent
staining with a polyclonal antibody raised against the Ad2 fiber. Lines
expressing the highest level of fiber protein were further characterized.

30 For the S8 cell complementing cell lines, to induce E1 expression, 0.3 µM
of dexamethasone was added to cell cultures 16-24 hours prior to challenge
with virus for optimal growth kinetics. For preparing viral plaques, 5×10^5

-76-

cells/well in 6 well plates are prepared and pre-induced with the same concentration of dexamethasone the day prior to infection with 0.5 μ M included at a final concentration in the agar overlay after infection.

F. Development of Cell Lines for Complementation of E1/E2a⁻ Vectors

5 The Adenovirus 5 genome was digested with Scal enzyme, separated on an agarose gel, and the 6,095 bp fragment containing the left end of the virus genome was isolated. The complete Adenovirus 5 genome is registered as Genbank accession #M73260, incorporated herein by reference, and the virus is available from the American Type Culture Collection, Manassas, Virginia, U.S.A.,
10 under accession number VR-5. The Scal 6,095 bp fragment was digested further with Clal at bp 917 and BglII at bp 3,328. The resulting 2,411 bp Clal to BglII fragment was purified from an agarose gel and ligated into the superlinker shuttle plasmid pSE280 (Invitrogen, San Diego, CA), which was digested with Clal and BglII, to form pSE280-E.

15 Polymerase chain reaction (PCR) was performed to synthesize DNA encoding an XhoI and Sall restriction site contiguous with Adenovirus 5 DNA bp 552 through 924. The primers which were employed were as follows:
5' end, Ad5 bp 552-585:

5'-GTCACTCGAGGACTCGGTC-GACTGAAAATGAGACATATTATCTGCCACGGA

20 CC-3' (SEQ ID No 36)

3' end, Ad5 bp 922-891:

5'-CGAGATCGATCACCTCCGGTACAAGGTTTGGCATAG-3' (SEQ ID No. 37)

 This amplified DNA fragment (sometimes hereinafter referred to as
Fragment A) then was digested with XhoI and Clal, which cleaves at the native
25 Clal site (bp 917), and ligated to the XhoI and Clal sites of pSE280-E, thus reconstituting the 5' end of the E1 region beginning 8 bp upstream of the ATG codon.

 PCR then was performed to amplify Adenovirus 5 DNA from bp 3,323
through 4,090 contiguous with an EcoRI restriction site. The primers which
30 were employed were as follows:

5' end, Ad5 bp 3323-3360:

-77-

5'-CATGAAGATCTGGAAGGTGCTGAGGTACGATGAGACC-3' (SEQ ID No. 38);

and

3' end, Ad5 bp 4090-4060:

5'-GCGACTTAAGCAGTCAGCTG-AGACAGCAAGACACTTGCTTGATCCAAATCC

5 -3' (SEQ ID No. 39).

This amplified DNA fragment (sometimes hereinafter referred to as Fragment B) was digested with BglII, thereby cutting at the Adenovirus 5 BglII site (bp 3,382) and EcoRI, and ligated to the BglII and EcoRI sites of pSE280-AE to reconstruct the complete E1a and E1b region from Adenovirus 5 bp 552 through 4,090. The resulting plasmid is designated pSE280-E1.

A construct containing the intact E1a/b region under the control of the synthetic promoter GRE5 was prepared as follows. The intact E1a/b region was excised from pSE280-E1, which was modified previously to contain a BamHI site 3' to the E1 gene, by digesting with XhoI and BamHI. The XhoI to BamHI fragment containing the E1a/b fragment was cloned into the unique XhoI and BamHI sites of pGRE5-2/EBV (U.S. Biochemicals, Cleveland, Ohio) to form pGRE5-E1).

Bacterial transformants containing the final construct were identified. Plasmid DNA was prepared and purified by banding in CsTFA prior to use for transfection of cells.

Construction of plasmid including Adenovirus 5 E2A sequence.

The Adenovirus 5 genome was digested with BamHI and SpeI, which cut at bp 21,562 and 27,080, respectively. Fragments were separated on an agarose gel and the 5,518 bp BamHI to SpeI fragment was isolated. The 5,518 bp BamHI to SpeI fragment was digested further with SmaI, which cuts at bp 23,912. The resulting 2,350 bp BamHI to SmaI fragment was purified from an agarose gel, and ligated into the superlinker shuttle plasmid pSE280, and digested with BamHI and SmaI to form pSE280-E2 BamHI-SmaI.

PCR then was performed to amplify Adenovirus 5 DNA from the SmaI site at bp 23,912 through 24,730 contiguous with NheI and EcoRI restriction sites.

The primers which were employed were as follows:

5' end, Ad5 bp 24,732-24,708:

-78-

5'-CACGAATTCGTCAGCGCTTCTCGTCGCGTCCAAGACCC-3' (SEQ ID No. 40);
3' end, Ad5 bp 23,912-23,934:

5'-CACCCCGGGGAGGCGGCGGCGACGGGGACGGG-3' (SEQ ID No. 41)

This amplified DNA fragment was digested with SmaI and EcoRI, and
5 ligated to the SmaI and EcoRI sites of pSE280-E2 Bam-Sma to reconstruct the
complete E2a region from Ad5 bp 24,730 through 21,562. The resulting
construct is pSE280-E2a.

In order to convert the BamHI site at the 3' end of E2a to a Sall site, the
E2a region was excised from pSE280-E2a by cutting with BamHI and NheI, and
10 recloned into the unique BamHI and NheI sites of pSE280. Subsequently, the
E2a region was excised from this construction with NheI and Sall in order to
clone into the NheI and Sall sites of the pMAMneo (Clonetech, Palo Alto, CA)
multiple cloning site in a 5' to 3' orientation, respectively. The resulting
construct is pMAMneo E2a.

15 Bacterial transformants containing the final pMAMneo-E2a were
identified. Plasmid DNA was prepared and purified by banding in CsTFA.
Circular plasmid DNA was linearized at the XmnI site within the ampicillin
resistance gene of pMAMneo-E2a, and further purified by the phenol/chloroform
extraction and ethanol precipitation prior to use for transfection of cells.

20 Transfection and selection of cells.

In general, this process involved the sequential introduction, by calcium
phosphate precipitation, or other means of DNA delivery, of two plasmid
constructions each with a different viral gene, into a single tissue culture cell.
The cells were transfected with a first construct and selected for expression of
25 the associated drug resistance gene to establish stable integrants. Individual cell
clones were established and assayed for function of the introduced viral gene.
Appropriate candidate clones then were transfected with a second construct
including a second viral gene and a second selectable marker. Transfected cells
then were selected to establish stable integrants of the second construct, and
30 cell clones were established. Cell clones were assayed for functional expression
of both viral genes.

-79-

A549 (ATCC Accession No. CCL-185) were used for transfection. Appropriate selection conditions were established for G418 and hygromycin B by standard kill curve determination.

Transfection of A549 cells with plasmids including E1 and E2a regions.

- 5 pMAMNeo-E2a was linearized with XmnI with the Amp^R gene, introduced into cells by transfection, and cells were selected for stable integration of this plasmid by G418 selection until drug resistant colonies arose. The clones were isolated and screened for E2a expression by staining for E2a protein with a polyclonal antiserum, and visualizing by immunofluorescence. E2a function was
- 10 screened by complementation of the temperature-sensitive mutant Ad5ts125 virus which contains a temperature-sensitive mutation in the E2a gene. (Van Der Vliet, et al., J. Virology, Vol. 15, pgs. 348-354 (1975)). Positive clones expressing the E2a gene were identified and used for transfection with the 7 kb EcoRV to XmnI fragment from pGRE5-E1, which contains the GRE5 promoted
- 15 E1a/b region plus the hygromycin^R gene. Cells were selected for hygromycin resistance and assayed for E1a/b expression by staining with a monoclonal antibody for the E1 protein (Oncogene Sciences, Uniondale, N.Y.). E1 function was assayed by ability to complement an E1-deleted vector. At this point, expression and function of E2a was verified as described above, thus
- 20 establishing the expression of E1a/b and E2a in the positive cell clones.

A transfected A549 (A549 (ATCC Accession No. CCL-185);) cell lines showed good E1a/b and E2a expression and was selected for further characterization. It was designated the S8 cell line.

25 G. Preparation of Adenoviral Vectors Containing Ad5. β gal. Δ F Genome in S8 Improved Fiber-Complementing Cell Lines

- To prepare adenoviral vectors containing Ad5. β gal. Δ F (Ad5. β gal. Δ F has been deposited the ATCC under accession number VR2636) in S8 cells containing alternative forms of TPL for enhancing the expression of fiber proteins, the protocol as described in Example 2 for preparing Ad5. β gal. Δ F in
- 30 211B cells was followed with the exception of pretreatment with 0.3 μ M dexamethasone for 24 hours as described above. Thus, viral particles with the wildtype Ad5 fiber protein on their surface and containing the fiberless

-80-

Ad5. β gal. Δ F genome were produced in 633 cells. Particles produced in 644 cells also contained the fiberless Ad5. β gal. Δ F genome, but had the chimeric 5T3H fiber protein, with the Ad3 fiber knob, on their surface.

Thus, these viral preparations, prepared as described herein, are useful for targeting delivery of the Ad5. β gal. Δ F, Ad5.GFP. Δ F, or other similarly constructed fiberless genome with either wild-type or modified fibers. Preferably for purposes herein the fibers are from an Ad serotype D virus, more preferably from Ad37.

EXAMPLE 6

10 Pseudotyping and Infectivity of Recombinant Adenoviral Vectors Produced with Improved Fiber-Complementing Cell Lines

A. Pseudotyping of Ad5. β gal. Δ F

To verify that adenoviral vectors were produced had altered tropisms, viral particles were purified from either 633 or 644 cells and were then Western
15 blotted and probed with a polyclonal rabbit antibody against the Ad2 fiber (which detects the Ad5 and chimeric 5T3H fiber proteins).

B. Infectivity of Cells with 633 or 644 Generated Virus Particles The cell lines, 633 or 644, prepared as described above, were infected with the indicated number of particles/cell of Ad5. β gal. Δ F and virus particles produced.
20 Virus was then used to infect selected cell lines, including 211B, MRC-5 human fibroblasts, A-10 rat aortic endothelial cells, and THP-1 human monocytic cells. Unbound virus was removed by washing the cells and the cells were further incubated at 37°C for 48 hours. Cells were then fixed with glutaraldehyde and stained with X-gal. The percentage of stained cells was then determined by light
25 microscopy where all experiments were done in triplicate.

The results indicated that adenoviral vectors could be retargeted by pseudotyping using packaging cell lines expressing different fiber proteins. Particles containing either fiber were equally infectious on 211B cells, while MRC-5 fibroblasts and THP-1 cells were more readily infected by virus containing
30 the chimeric fiber. The A-10 rat endothelial cells were more readily infected by particles containing the wildtype Ad5 fiber protein.

-81-

EXAMPLE 7

Transient Transcomplementation

The ability of adenovirus type 5 (Ad5) to deliver therapeutic genes to cells is mediated by the interaction of the adenoviral fiber protein with the coxsackievirus-adenoviral receptor (CAR). Because a wide-range of cells express CAR, it was thought that it would be difficult to use adenoviruses to deliver genes to specific cell types. A system for testing modified fiber genes to identify tropisms of interest is described in copending U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265 on January 14, 2000). An *in vitro* system has been developed that involves infection of tissue culture cells with a fiber-deleted Ad and transient co-transfection with a plasmid directing fiber expression. This system allows one to produce and evaluate modified fibers expressed on a viral particle. This system can be used to produce therapeutic quantities of adenoviral vectors with modified fiber proteins, with such fibers having a new tropism added by insertion of a desired ligand into the fiber gene. These fibers may also have the natural tropism (*i.e.* binding to CAR) ablated.

Plasmids used were pDV60 and pDV55 were prepared as described herein and in U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265 on January 14, 2000). pDV60 is an pcDNA3.1-based expression plasmid that contains the CMV promoter, Ad5 tripartite leader, an intron, and the Ad5 fiber gene sequence. pDV55 contains no fiber gene and serves as the negative control. Ad5. β gal. Δ F and 211B are described above. 293T cells are identical to 293 cells except they express an integrated SV40 large T antigen gene. HDF cells are human diploid fibroblasts. 293T cells express CAR and α_v integrins; HDF cells express α_v integrins but no CAR. Transfections with fiber expression plasmids were performed with Lipofectamine (GIBCO-BRL) using 20mg DNA and 50ml Lipofectamine per 15cm dish. Cells were maintained in DMEM supplemented with 10% fetal bovine serum.

The fiber deletion mutation of Ad5. β gal. Δ F is complemented in *trans* by passaging virions through 211B, a cell line that stably expresses functional Ad5.

-82-

fiber. The present system was designed to complement Ad5. β gal. Δ F by modified fibers expressed from transfected episomal plasmids in 293T cells. The result is a simplified and rapid method to incorporate modified fibers on a viral particle containing the Ad5. β gal. Δ F genome that does not require propagation of the virus.

The feasibility of transcomplementation of Ad5. β gal. Δ F with episomal fiber-expressing plasmids was demonstrated in the following experiment. 293T cells were transfected with one of two plasmids: pDV55, which expresses no fiber or pDV60, which expresses wildtype Ad5 fiber. Fiber expression persists for at least six days. Twenty-four hours after transfection, these cells were infected at 2000 particles/cell with Ad5. β gal. Δ F passaged through 211B cells. Seventy-two hours later, a crude viral lysate (CVL) was generated by exposing the cells to five freeze-thaw cycles. Viral particles were purified by cesium chloride gradient centrifugation. The resulting virions incorporated the fiber expressed from the episomal plasmid, as confirmed by Western blots performed with an antibody specific to the Ad5 fiber.

Episomal plasmid transcomplementation system is suitable for quickly expressing and evaluating the properties of modified fibers in the context of a viral particle. Episomal plasmid transcomplementation will also be of great utility for quickly evaluating a bank of modified fibers for other binding properties, including new tropisms and the ablation of the native tropism. In addition to the rapid generation and testing of large numbers of modified fibers, there are other advantages to the Ad5. β gal. Δ F transcomplementation system in terms of production and safety. Episomal plasmid transcomplementation has the inherent advantage over transcomplementation in that it is not necessary to make a stable cell line for every modified fiber for complementation with Ad5. β gal. Δ F. Because the Ad5. β gal. Δ F is deleted in E1, E3 and fiber, there is an additional gene deletion, which should render it very suitable for gene therapy. In addition, the presence of the fiber gene deletion decreases the opportunity to generate replication-competent virus via recombination in the packaging cells. A single Ad vector preparation can be retargeted to any number of different cell types simply by transfecting the cells with the appropriate fiber-expression construct.

-83-

EXAMPLE 8

Preparation of Adenoviral Gene Delivery Vectors Containing the Ad37 fiber protein

This example describes construction of packaging cell lines expressing the Ad37 fiber protein, and their use in generating particles of a fiber-deleted Ad vector (such as Ad5. β gal. Δ F) containing this fiber protein. The fiber protein is attached to the viral capsid by binding to the penton base protein through its N-terminus, and the Ad37 fiber was modified in order to make its N-terminal sequence more closely match that of the Ad5 protein to ensure that it would efficiently bind the Ad5 penton base in these vectors.

A. Materials and methods

Cell lines and wild-type adenovirus. Human A549 lung carcinoma epithelial cells and human Chang C conjunctival cells (American Type Culture Collection) were maintained in complete Dulbecco's Modified Eagle Medium (DMEM) with 10% fetal bovine serum. Wild-type Ad19p and Ad37 (ATCC) were propagated in A549 cells and purified by banding on CsCl₂ density gradients as previously described (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). Viral protein concentration was determined by the Bio-Rad Protein Assay, and was used to calculate the number of viral particles based on the known molecular weight of Ad2 virions ($1 \mu\text{g} = 4 \times 10^9$ particles).

B. Construction of the Ad37 fiber expressing cell lines and the recombinant Ad37 knob protein.**1. Construction of an Expression Plasmid for the Ad37 Fiber Protein (pDV80)**

The plasmid designated pDV80 (see, SEQ ID No. 42) prepared for expression of the Ad37 fiber protein in mammalian cells, uses the same regulatory elements as the elements in pDV60, pDV67, and pDV69 to express the Ad37 fiber in packaging lines. It was constructed in two steps.

First, the Ad37 fiber open reading frame was amplified from Ad37 genomic DNA using synthetic oligonucleotide primers, L37: 5' TGT CCT GGA TCC AAG ATG AAG CGC GCC CGC CCC AGC GAA GAT GAC TTC 3' (SEQ ID NO. 43) and 37FR: 5' AAA CAC GGC GGC CGC TCT TTC ATT CTT G 3' (SEQ ID NO. 44). L37 contains nucleotides (underlined) that differ from the Ad37

-84-

genomic sequence in order to add a unique *Bam*H1 site (bold) before the start codon (italicized) and to create point mutations that make the N-terminal sequence of the fiber more closely match the N-terminal sequence of the Ad5 fiber protein as follows:

5 Ad37 MSKRLRVEDDFNPVYPY (SEQ ID No. 45)
 ↓↓↓↓↓↓
 KRARPS (SEQ ID No. 46)
 Ad5 MKRARPS EDTFNPVYPY (SEQ ID No. 47).

37FR also incorporates a unique *Not*1 site (bold). The PCR product was inserted
 10 into the *Bam*H1 and *Not*1 sites of pCDNA3.1zeo(+) (Invitrogen) to create pDV78. The correct sequence of the Ad37 fiber protein, including inserted changes, was confirmed by sequencing.

Two point mutations in the fiber gene in the 705 line, S356 to P356 and
 1362 to T362, were discovered by the sequencing. The mutations are not in the
 15 receptor binding domain in Ad37 fiber gene in the 705 cell line. They are buried in the knob trimer interface. To confirm that these mutations do not affect receptor binding, the Ad37 fiber protein with the correct sequence was recloned, and 293T cells transfected with the virus and subsequently infected with
 Ad5.GFPΔF to produce Ad37 pseudotyped virus. The results were the same as
 20 the results of the experiments with Ad37 pseudotyped virus produced from line 705 (see, Wu *et al.* (2001) *Virology* 279:78-89).

Second, a 1.2 kb *Bam* H1/*Bgl* II fragment containing an adenovirus type 5 tripartite leader was excised from pDV55 (see EXAMPLE 3) and inserted into the
Bam H1 site of pDV78 to create pDV80 (SEQ ID No 42). Plasmid pDV80 has
 25 been deposited in the ATCC under accession number PTA-1147.

2. Construction of the recombinant Ad37 knob protein

Recombinant Ad37 knob protein containing an N-terminal T7•Tag was produced in *E. coli* using the PET expression system (Novagen). Ad37 fiber DNA (GenBank accession number U69132) was PCR amplified from wild-type Ad37
 30 genomic DNA using the following primers (SEQ ID Nos. 48 and 49):
 5' GGATCCATGGGATACTTGGTAGCA 3' (*Bam*H1 site underlined and
 5' GCAACTCGAGTCATTCTTGGGCAATATAGG 3' (*Xho*I site underlined).

-85-

The PCR reactions were performed at 94°C (denaturation), 55°C (annealing), 72°C (extension, 30 cycles) using *Taq* DNA polymerase (Qiagen). The amplified DNA fragments, which contained residues 172 to 365 of the Ad37 fiber protein with the addition of an N-terminal start codon (italicized), were purified and subcloned into the pCR-TOPO vector using the TA-Cloning Kit (Invitrogen). No replication errors were found by DNA sequencing. Plasmids from cultured transformed colonies were purified and digested with *Bam*HI and *Xho*I. The fragment was inserted into the *Bam*HI and *Xho*I sites of the bacterial expression vector, pET21a (Novagen), and transformed into (DE3)pLYS S expression cells (Invitrogen). Colonies were selected for knob expression by induction with 1 mM IPTG for four hours at 37°C and knob expression was determined by SDS-PAGE. The colony displaying highest knob expression was used for large-scale knob expression and induced with 0.5 mM IPTG at 30°C for four hours.

The recombinant T7•Tagged Ad37 knob protein was purified from sonicated bacteria using the T7•Tag Affinity Purification Kit as recommended by the manufacturer (Novagen). Recovered protein was analyzed for purity by SDS-PAGE followed by Coomassie staining or Western blotting with an HRP-conjugated α -T7•Tag monoclonal antibody as described by the manufacturer (Novagen) or an α -Ad37 fiber rabbit antibody.

3. Preparation of Cell Lines that Express the Ad37 fiber protein

Plasmid pDV80 DNA was purified using the Qiagen method and electroporated into the adenovirus-complementing cell line E1-2a S8 (see Examples herein; see also, Gorziglia *et al.* (1996) *J. Virology* 70:4173-4178; and Von Seggern *et al.* (1998) *J. Gen. Virol.* 79:1461-1468). Stable clones were selected with 600 μ g/ml zeocin (Invitrogen).

Clones were expanded and were screened for fiber expression by indirect immunofluorescence (Von Seggern *et al.* (1998) *J. Gen. Virol.* 79:1461-1468) using a rabbit polyclonal antibody directed against the Ad37 fiber (α -Ad37 fiber rabbit antibody) raised by immunizing rabbits with recombinant Ad37 fiber protein. Two clones (lines 705 and 731) that expressed the protein at a uniformly high level were selected.

-86-

EXAMPLE 9

Production of Pseudotyped Ad Vector Particles

To generate vector particles equipped ('pseudotyped') with the Ad37 fiber protein, the Ad37 fiber-expressing 705 cells were infected (approximately 5 1000 particles/cell) with Ad5. β gal. Δ F or with Ad5.GFP. Δ F.

Materials and methods

Ad5. β gal. Δ F

The construction of Ad5. β gal. Δ F is described in Example 2 (it has been deposited on January 15, 1999, with the ATCC as listed above under accession 10 number VR2636; see also, Von Seggern *et al.* (1999) *J. Virol.* 73:1601-1608; copending U.S. application Serial No. 09/482,682 filed January 14, 2000, and also International PCT application No. PCT/US00/00265, filed January 14, 2000).

Ad5.GFP. Δ F

15 Ad5.GFP. Δ F was constructed by recombination in bacteria using a modification of the AdEasy System (see, U.S. Patent No. 5,922,576; see, also He *et al.* (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:2509-2514; the system is publicly available from the authors and other sources).

First, a fiber-deleted genomic plasmid was constructed by removing the 20 fiber gene from pAdEasy-1 (see, U.S. Patent No. 5,922,576; and He *et al.* (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:2509-2514; the AdEasy system and vectors are publicly available from He *et al.* at Johns Hopkins University). Plasmid pAdEasy-1 contains the entire Ad5 genome, except for nucleotides 1-3,533, which encompass the E1 genes, and nucleotides 28,130-30,820, which 25 encompass the E3 gene.

Plasmid pDV43 (see Example 2; see, also Von Seggern *et al.* (1999) *J. Virol.* 73:1601-1608) was digested with *Pac*1, the ends blunted by treatment with the large fragment of *E. coli* DNA polymerase and dNTPs, and the product re-ligated to produce plasmid pDV76. The resulting plasmid pDV76 is identical to 30 pDV43 except for loss of the *Pac*1 site and contains the right end of the Ad5 genome with E3 and fiber deletions. A 4.23 kb fragment from pDV76 was amplified using the oligonucleotide primers (SEQ ID Nos. 50 and 51:

-87-

5' CGC GCT GAC TCT TAA GGA CTA GTT TC 3', including the unique *Spe*1 site in the Ad5 genome (bold); and 5' GCG CTT AAT TAA CAT CAT CAA TAA TAT ACC TTA TTT T 3', including a new *Pac*1 site (bold) adjacent to the right Ad5 ITR. Hence the resulting PCR amplified fragment contains nucleotides
5 27,082 to 35,935 of the Ad5 genome with deletions of nucleotides 28,133 to 32,743 (the E3 and fiber genes), and was used to replace the corresponding *Spe*1/*Pac*1 fragment of pAdEasy 1 (see, U.S. Patent No. 5,922,576) to create pDV77.

Second, *E. coli* strain BJ5183 was electroporated with a mixture of
10 pDV77 and *Pme*1-linearized pAdTrack as described (U.S. Patent No. 5,922,576; He *et al.* (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:2509-2514), and DNA was isolated from kanamycin-resistant colonies. The resulting plasmid, pDV83, contains a complete Ad5 genome with E1-, E3-, and fiber-deletions with a CMV-driven GFP reporter gene inserted at the site of the E1 deletion. The full length
15 Ad chromosome was isolated by *Pac*1 digestion, and transfected into the E1- and fiber-complementing 633 cells (Von Seggern *et al.* (2000) *J. Virol.* 74:354-362). The 633 cells were produced by electroporating pDV67 (SEQ ID No. 30, deposited under ATCC accession number PTA-1145) into the E1-2a S8 cells, described above. The recovered virus Ad5.GFP.ΔF was then plaque purified by
20 plating on 633 cells and virus stocks were prepared by freeze-thawing cell pellets.

Ad5-pseudotype particle production

Particles with Ad5 fiber

Ad5-pseudotyped particles were generated by virus growth in 633 cells,
25 which express the wild type Ad5 fiber protein. Viral particles were isolated and purified over CsCl gradients (Von Seggern *et al.* (1999) *J. Virol.* 73:1601-1608; purified by centrifugation on preformed 15-40% CsCl gradients (111,000 x g for three hours at 4°C)). For analysis of viral proteins, ten μg of the purified particles were electrophoresed on 8-16% gradient gels and the protein
30 transferred to nylon membranes. The resulting blot was probed with rabbit polyclonal antibodies raised against recombinant Ad37 fiber or Ad5 fiber or penton base proteins expressed in baculovirus-infected cells.

-88-

Particles with Ad37 fiber

Cells from the Ad37 fiber producing cell line 705 were infected at approximately 1000 particles/cell with Ad5. β gal. Δ F or with Ad5.GFP. Δ F. Viral particles were isolated and purified over CsCl gradients. The bands were
5 harvested, dialyzed into storage buffer (10 mM Tris-pH 8.1, 0.9% NaCl, and 10% glycerol), aliquoted and stored at -70°C.

Viral protein analyses

For analysis of viral proteins, 10 μ g of purified Ad5. β gal. Δ F particles with no fiber (grown in 293 cells), the Ad5 fiber (grown in 633 cells), or the Ad37
10 fiber (grown in 705 cells) were electrophoresed by 8-16% polyacrylamide gradient SDS-PAGE and the proteins were transferred to nylon membranes. The blot was then probed with α -Ad37 fiber rabbit antibody. Ad5 fiber and penton base were detected by reprobing the blot with polyclonal antibodies raised against recombinant proteins expressed in baculovirus-infected cells (Wickman *et*
15 *al.* (1993) *Cell* 2:309-319).

Adenovirus infection and cell binding assays

Adherent Chang C and A549 cells were infected with GFP expressing Ad5 vectors containing the Ad5 fiber (Ad5.GFP. Δ F/5F) or the Ad37 fiber (Ad5.GFP. Δ F/37F) at 10,000 particles per cell for 3 hours at 37°C, 5% CO₂ in
20 DMEM, 10% FCS. Cells were washed twice with saline and then cultured overnight at 37°C, 5% CO₂. The next day, the cells were detached with buffer containing 0.05% (w/v) trypsin and 0.5 mM EDTA (Boehringer Mannheim) for 5 minutes at 37°C. Suspended cells were washed once with PBS and then resuspended in phosphate-buffered saline (PBS), pH 7.4. GFP fluorescence was
25 measured with a FACScan flow cytometer. A threshold established by the fluorescence of uninfected cells was used to distinguish cells expressing GFP. To assess the role of CAR in Ad infection, 10,000 attached cells were pre-incubated with 180 μ g/ml RmcB, a function-blocking anti-CAR monoclonal antibody (Hsu *et al.* (1988) *J. Virol.* 62:1647-1652), in complete DMEM for 1
30 hour at 4°C. A small volume containing Ad5.GFP. Δ F/5F or Ad5.GFP. Δ F/37F was then added at 10,000 particles per cell. The cells were infected for 3 hours, cultured overnight, harvested, and analyzed for GFP expression. Percent

-89-

cells expressing GFP was determined by the percent of cells detected above a threshold set by the fluorescence of uninfected Chang C cells.

To measure adenovirus binding to cells, wild type Ad37 was labeled with ^{125}I using Iodogen (Pierce) according to manufacturer instructions and separated from free ^{125}I by gel filtration as described (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). Binding of radiolabeled wild type Ad37 on Chang C cells was then quantitated as described (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). Non-specific binding was determined by incubating cells and labeled Ad37 particles in the presence of 100-fold concentration of unlabeled Ad37. Specific binding was calculated by subtracting the non-specific binding from the total cpm bound. To examine if divalent cations are required for binding, 10 mM ethylenediaminetetraacetic acid (EDTA) or various concentrations of CaCl_2 , or MgCl_2 were added to cells before incubation with labeled virus. To examine if the receptor for Ad37 is a protein, cells were pretreated with 10 $\mu\text{g}/\text{ml}$ trypsin (GIBCO), subtilisin (Sigma), proteinase K (Boehringer-Mannheim), and bromelain (Sigma) at 37 °C for 1 hour, then washed twice with complete DMEM before adding labeled virus. Cells were >95% viable after protease treatment.

Ad37 binding to conjunctival cells is calcium-dependent. Specific ^{125}I -labeled Ad37 binding to Chang C cells was measured in the presence of 10 mM EDTA and in the presence of varying concentrations of calcium chloride or magnesium chloride. Specific binding was determined by subtracting the nonspecific counts in the presence of 100-fold excess unlabeled virus from the total counts.

Pretreatment of conjunctival cells with proteases inhibits Ad37 binding. Change C cells were pretreated with various proteases for 1 hour before binding ^{125}I -labeled Ad37 to the cells. Nonspecific binding was measured by adding 100-fold unlabeled Ad37 to cells with ^{125}I -labeled Ad37 and subtracting from total counts for specific binding. Percent inhibition represents the difference in specific binding of untreated cells and pretreated cells as a percentage of the specific binding of untreated cells.

-90-

Virus overlay protein blot assay (VOPBA)

- For VOPBA of human conjunctival membrane proteins probed with Ad37 in the presence of EDTA or calcium chloride, Chang C membrane fractions were separated by 8% SDS-PAGE and transferred to a PDVF membrane. The
- 5 membrane was subsequently probed with or without whole Ad37 particles, a polyclonal antibody against Ad37 fiber, and finally a horseradish peroxidase conjugated anti-rabbit antibody, in the presence of EDTA or calcium chloride. Transferred Chang C membrane proteins were probed with recombinant Ad37 knob protein, instead of Ad37 knob, in the presence of calcium chloride.
- 10 Confluent monolayers of Chang C and A549 cells were detached by scraping, pelleted by centrifugation, and then resuspended in 250 mM sucrose, 20 mM HEPES, pH 7.0, 1 mM EDTA, and 2 μ g/ml aprotinin and leupeptin. Cells were transferred into a dounce homogenizer and disrupted with 30 strokes. Organelles and nuclei were pelleted at 500g for 15 min. Plasma membrane
- 15 fragments were then pelleted from the supernatant of cell lysates at 200,000g for 1 hour and then resuspended in 10 mM Tris•Cl, pH 8.1, 10 μ g/ml aprotinin and leupeptin.
- Cell membranes of Chang C or A549 cells were incubated (1:1) with a 2% SDS, non-reducing buffer and separated on an 8% polyacrylamide gel
- 20 without boiling. Membrane proteins were then electroblotted onto a PVDF membrane (Immobilon-P) and blocked in 5% (w/v) milk in PBS, pH 7.4, 0.02% Tween-20 (PBS-T). After blocking, the membrane was incubated with 1 μ g/ml wild-type Ad19p or Ad37 in 0.5% (w/v) milk in PBS-T, 1 mM CaCl₂, for 1 hour at room temperature. The membrane was then washed once with phosphate-
- 25 buffered saline, pH 7.4 (PBS), 1 mM CaCl₂, and incubated with 1:500 dilution of α -Ad37 fiber rabbit antibody in 0.5% (w/v) milk in PBS-T, 1 mM CaCl₂, for 30 minutes at room temperature. The membrane was washed again with PBS, 1 mM CaCl₂, and incubated with 1:5000 dilution of horseradish peroxidase (HRP) conjugated α -rabbit antibody (Sigma) in 0.5% (w/v) milk in PBS-T, 1 mM CaCl₂,
- 30 for 30 minutes at room temperature. The membrane was washed four times in PBS, 1 mM CaCl₂, once with PBS-T, 1 mM CaCl₂, and once in 1 mM CaCl₂. The blot was developed with enhanced chemiluminescence reagents (Pierce) for 5

-91-

minutes and placed onto a piece of Biomax film (Kodak) for 5 seconds to 1 minute. For divalent metal cation experiments, membranes were incubated in the presence of 2 mM EDTA instead of 1 mM CaCl_2 in all solutions. To assay fiber knob binding to cell membrane proteins, membrane filters were incubated with 1 $\mu\text{g/ml}$ purified T7-tagged Ad37 knob protein in Tris-buffered saline, 0.1% Tween-20, 1 mM CaCl_2 , for 1 hour at room temperature. α -Ad37 fiber rabbit antibody and HRP-conjugated anti-rabbit antibody were applied and the membrane was developed with substrate solution as described above.

Results: Comparison of adenovirus infection of human conjunctival and lung epithelial cells with virus particles retargeted with Ad5 or Ad37 fiber proteins

Packaging cell lines producing the Ad37 fiber protein were generated. Since the N-terminal amino acid sequences of the Ad5 and Ad37 fiber proteins differ significantly, and to ensure that the Ad37 fiber would be efficiently incorporated into Ad5 vector particles, several residues in the wild-type Ad37 fiber were mutated to more closely match the Ad5 sequence. Stable cell lines producing this fiber under control of the CMV promoter and the adenovirus type 5 tripartite leader were then generated and screened for fiber expression by indirect immunofluorescence. One clone (line 705), which expressed the Ad37 fiber at a high level, was selected for further study.

Cells from one cell line 633, which expresses the wild-type Ad5 fiber protein, and line 705 were infected with a fiber-deleted Ad5 vector carrying a β galactosidase reporter gene. The resulting vector particles contained the Ad5 fiber protein (Ad5. β gal. Δ F/5F) and the Ad37 fiber protein (Ad5. β gal. Δ F/37F), respectively. Incorporation of the correct fiber protein into viral particles was verified by Western blotting. Adenoviral vectors containing the GFP reporter gene, Ad5.GFP. Δ F/5F and Ad5.GFP. Δ F/37F, were created in the same fashion.

Infection of a variety of cell types using the retargeted adenovirus particles was examined. As assayed by GFP fluorescence, Ad5.GFP. Δ F/5F exhibited good gene delivery to lung epithelial (A549) and conjunctival cells (Chang C). In contrast, Ad5.GFP. Δ F/37F efficiently delivered GFP to Chang C cells, but exhibited very poor gene delivery to A549 cells. Although CAR is expressed on the surface of A549 cells, as indicated by Ad5.GFP. Δ F/5F

-92-

infection, Ad5.GFP. Δ F/37F was unable to infect these cells efficiently. This experiment shows that the Ad37 fiber protein can confer preferential infection of human conjunctival cells, but not CAR-expressing human lung epithelial cells.

Hence CAR is not the primary receptor for Ad37. Recent studies reported
5 that expression of CAR on the surface of chinese hamster ovary (CHO) cells did not improve Ad37 binding (Arnberg *et al.* (2000) *J. Virol.* 74:42-48), implying that Ad37 does not use CAR as a primary receptor. In order to verify this on human conjunctival cells, A549 and Chang C cells were pretreated with RmcB (Hsu *et al.* (1988) *J. Virol.* 62:1647-1652), a function-blocking monoclonal
10 antibody against CAR. The RmcB antibody inhibited infection of A549 cells by Ad5.GFP. Δ F/5F, but it had little effect on infection of Chang C cells by Ad5.GFP. Δ F/37F. This indicates that CAR is not the primary receptor for Ad37 on Chang C conjunctival cells.

Ad37 binding to conjunctival cells requires divalent metal cations. It has
15 been proposed (Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915) that a combination of fiber binding to CAR and penton base binding to α_v -integrins allows some adenovirus serotypes to attach to cells. Although α_v -integrin binding to the RGD motif of the adenovirus penton base is of relatively low affinity (Wickman *et al.* (1993) *Cell* 2:309-319), it may nonetheless contribute to viral attachment to the
20 cell surface. Ad37 shows a particularly strong affinity for binding to integrin $\alpha_v\beta_5$ (Mathias *et al.* (1998) *J. Virol.* 72:8669-8675), suggesting that integrin $\alpha_v\beta_5$ might be a primary receptor for Ad37. Binding of the RGD motif by α_v -integrins requires the presence of divalent cations, such as calcium or magnesium (Stuiver *et al.* (1996) *J. Cell Physiol.* 168:521-531). In contrast, no divalent cations were
25 required for binding in the CAR-Ad12 knob complex (Bewley *et al.* (1999) *Science* 286:1579-1583).

To investigate the potential role of α_v -integrins and divalent metal cations in Ad37 receptor binding, 125 I-labeled Ad37 binding to Chang C cells was examined in the absence or presence of EDTA. EDTA inhibited Ad37 binding to
30 conjunctival cells but did not alter Ad5 binding. These findings suggest a requirement for divalent metals for Ad37 binding.

-93-

The presence of either calcium or magnesium ions helps $\alpha_v\beta_5$ organize in focal contacts (Stuiver *et al.* (1996) *J. Cell Physiol.* 168:521-531), suggesting that calcium and magnesium aid in integrin $\alpha_v\beta_5$ function. To further test the potential role of integrin $\alpha_v\beta_5$ in Ad37 cell attachment, ^{125}I -labeled Ad37 binding to Chang C cells was measured in the presence of varying concentrations of calcium or magnesium chloride. Magnesium ions had little effect on Ad37 binding to Chang C cells. In contrast, calcium ions dramatically enhanced Ad37 binding to Chang C cells. The optimal concentration of calcium chloride for Ad37 binding was 1 mM, while higher concentrations of calcium actually decreased virus binding to cells. The fact that calcium, but not magnesium, promoted Ad37 attachment is not consistent with integrin $\alpha_v\beta_5$ as the primary receptor for viral attachment to the cells since either metal will support ligand binding to integrin $\alpha_v\beta_5$. Moreover, A549 cells express abundant α_v -integrins (Mathias *et al.* (1998) *J. Virol.* 72:8669-8675) but were unable to support efficient binding of Ad37.

Wild-type Ad37 particles bind to three conjunctival membrane proteins. Recent studies reported that protease treatment of CHO cells abolished Ad37 binding (Arnberg *et al.* (2000) *J. Virol.* 74:42-48), implying that Ad37 bound to a protein receptor on CHO cells. Scatchard analysis of Ad37 binding to Chang C cells showed that each cell expresses approximately 24,000 fiber binding sites (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). To determine if the Ad37 binding site on human conjunctival cells is also a protein, Chang C cells were treated with different proteases prior to measuring binding of ^{125}I -labeled Ad37. Digestion of surface proteins by all four proteases inhibited Ad37 binding to Chang C cells by greater than 50%. This finding showed that Ad37 also binds to a protein receptor on Chang C cells.

Virus overlay protein blot assays (VOPBAs) were used to identify candidate viral protein receptors. This Western blot technique uses intact viral particles in place of antibodies to probe viral-receptors interactions. VOPBA was used herein to identify Chang C membrane proteins that bind to Ad37. In the absence of Ad37 particles, no protein bands were observed, while addition of virus in the absence of calcium revealed binding to a single 45 kDa protein. In

-94-

the presence of 1 mM calcium chloride, Ad37 reacted with three proteins with approximate molecular weights of 45, 50 and 60 kDa. The same three proteins were detected using a recombinant Ad37 fiber knob alone, indicating that Ad37-receptor interactions are fiber mediated and do not involve interactions of other capsid proteins such as the penton base. The size of the calcium-independent protein (45 kDa) is very similar to the known molecular weight of CAR. A direct comparison of the Ad37 VOPBA and a CAR Western blot showed that the 45 kDa receptor co-migrates with CAR on SDS-PAGE. Moreover, two other members of subgroup D adenoviruses, Ad9 and AD15, have been shown to bind to CAR (Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915).

Since CAR does not appear to mediate Ad37 binding on intact Chang C cells, the possibility that the 50 or 60 kDa protein serves this function was tested by examining an adenovirus serotype that does not bind to Chang C cells. Ad19p, a closely related subgroup D adenovirus, binds poorly to Chang C cells (Huang *et al.* (1999) *J. Virol.* 73:2798-2802) and Ad19p recognition of the Ad37 receptor is therefore unlikely. Ad19p particles bound to the 45 and 60 kDa receptors in the VOPBA, but did not bind to the 50 kDa receptor. Moreover, the 50 kDa receptor is expressed on Chang C cells, but not A549 cells, which only support low levels of Ad37 binding and infection. Taken together, these data indicate that the 50 kDa protein is a primary candidate receptor for Ad37 on human conjunctival cells.

Discussion

The identification of the CAR protein as a major adenovirus receptor does not explain why certain subgroup D members, such as Ad37, preferentially infect ocular cells. A 50 kDa human conjunctival cell membrane protein is identified herein as a primary candidate for the receptor for Ad37. This 50 kDa protein is not present on A549 lung epithelial cells. Ad37 binding to this receptor is calcium-dependent, which is consistent with Ad37 binding and infection experiments. Ad37 also bound to a 60 kDa protein that is present on human conjunctival and lung epithelial cells. It does not, however, appear to be serotype specific. The molecular weights of MHC class I heavy chain, which has been proposed as a receptor for Ad5, and $\alpha_v\beta_3$ and $\alpha_v\beta_5$ integrins, receptors for

-95-

the penton base, are distinct from the 50 or to kDa receptor characterized in this study.

The studies of Ad37-receptor interaction using VOPBAs are consistent with previous studies showing that subgroup D adenoviruses can bind to the extracellular domain of CAR (Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915). Biochemical and structural studies on knob-CAR interactions indicate that the CAR binding site is located on the AB-loop of the fiber knob. Alignment of the fiber sequences of Ad37 and other adenoviruses reveals that the AB-loop of Ad37 is similar to those of Ad12 and Ad5. Moreover, a phylogenetic tree of adenovirus knobs (Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915) shows that fiber proteins of subgroup D are similar to those of subgroup C and E, which use CAR as their primary receptor. Ad37 does not, however, appear to effectively use CAR as a primary receptor, as demonstrated by virus binding and infection studies on Chang C conjunctival cells and A549 lung epithelial cells.

It has been reported that Ad37 uses sialic acid as a receptor on chinese hamster ovary (CHO) cells and human lung carcinoma (A549) cells (Arnberg *et al.* ((2000) *J. Virol.* 74:42-48). Human conjunctival cells were not studied. Human corneal epithelial (HCE) cells were the only ocular cell line studied and Ad37 binds relatively poorly to these cells, compared to binding on A549 cells (Arnberg *et al.* ((2000) *J. Virol.* 74:42-48). In addition, $8,4 \times 10^7$ wheat germ agglutinin molecules per cell were required to significantly inhibit Ad37 binding to sialic acid on sialic acid positive CHO cells (Arnberg *et al.* (2000) *J. Virol.* 74:42-48), three orders of magnitude higher than the number of Ad37 receptors on Chang C conjunctival cells (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). Clearly, sialic acid is not the only factor responsible for Ad37 binding to the cell surface and its influence on Ad37 tropism is unclear.

The results herein show that Ad37 selects a 50 kDa cellular receptor for binding to conjunctival cells, but it is possible that sialic acid also plays a role in this interaction. The characterization and identification of the Ad37 receptor have therapeutic implications and also explain the different tropism of Ad37. The 50 kDa receptor for Ad37 may also be the receptor for other subgroup D adenoviruses that cause severe cases of EKC, Ad19a and Ad8. Ad19p is a

-96-

nonpathogenic variant of Ad19 (Arnberg *et al.* (1998) *Virology* 227:239-244) while Ad19A, along with Ad8 and Ad37, are major causes of EKC. Ad19a and Ad37 have identical fiber proteins (Arnberg *et al.* (1998) *Virology* 227:239-244) and have similar tropism in vivo. Ad8, Ad19a, and Ad37 agglutinate dog and guinea pig erythrocytes more effectively than four other serotypes that are associated with less severe forms of conjunctivitis (Arnberg *et al.* (1998) *Virology* 227:239-244), implying that the receptors of Ad18, Ad19A, and Ad37 have similar characteristics. Hence, this 50 kDa receptor is an attractive drug target against EKC caused by adenoviruses to provide therapeutic intervention of ocular diseases associated with these viruses.

EXAMPLE 10

Targeting of the Ad5 vector to photoreceptor cells

The fiber-deleted adenovirus vector Ad5.GFP. Δ F was propagated in 705 cells, which express a modified Ad37 fiber protein. Viral particles (Ad5.GFP. Δ f/37F) were harvested, CsCl-purified and dialized into 0.9% NaCl, 10 mM Tris, pH 8.1, and 10% glycerol. Two to three μ l of the resulting solution, containing approximately 1×10^9 particles/ μ l was injected into the vitreous chamber of a mouse eye. Seven days post-injection, eyes were harvested, fixed with paraformaldehyde and cryo-sectioned. Sections were stained with an anti-rhodopsin antibody to identify photoreceptor cells and with DAPI to show all cell nuclei. The resulting sections showed red anti-rhodopsin staining in the photoreceptors, blue DAPI-stained nuclei, and green GFP staining in any transduced cells. The results revealed substantially exclusive transduction of photoreceptors. Co-localization of rhodopsin staining and GFP expression indicated selective transduction of photoreceptor cells.

As a control, contralateral eyes were injected with a stock of the fiber-deleted vector AD5. β gal. Δ F grown in the same Ad37 fiber-expressing cells. Since this virus (Ad5. β gal. Δ F/37F) produces β gal rather than GFP, the green staining is absent from the photoreceptors.

Additional experiments using the AD37 fiber for targeting to the photoreceptor cells have been performed. Subretinal and intravitreal injection have been used in mouse models and the results demonstrate targeting to the

-97-

photoreceptors. As with intravitreally injected eyes, the major cell type infected via subretinal administration was the photoreceptor.

- As noted, Ad5.GFP. Δ F /37F infected Chang C cells efficiently, but A549 cells poorly. Ad37 fiber protein confers preferential infection on human
- 5 conjunctival cells, but not CAR-expressing human lung epithelial cells. Binding to conjunctival cells requires divalent cations.

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims:

-98-

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule, comprising:
adenovirus inverted terminal repeat sequences; an adenovirus packaging
signal operatively linked thereto; and a photoreceptor-specific promoter.
- 5 2. The isolated nucleic acid molecule of claim 1, further comprising a
nucleic acid encoding a therapeutic product operatively linked to the promoter.
3. The isolated nucleic acid molecule of claim 1, wherein the
promoter is a rhodopsin promoter.
4. The nucleic acid molecule of claim 1, wherein the adenovirus
10 genome does not encode a functional fiber protein such that packaging the
nucleic acid requires complementation in a packaging cell.
5. A recombinant adenovirus vector, comprising the nucleic acid
molecule of any of claims 1-4 packaged therein.
6. A recombinant adenovirus vector of claim 5, wherein inverted
15 terminal repeat sequences (ITR) and a packaging signal are derived from
adenovirus type 2 or adenovirus type 5.
7. A recombinant adenovirus vector of claim 5, wherein the virus
comprises a fiber protein.
8. A recombinant adenovirus vector of claim 7, wherein the fiber
20 protein selectively binds to photoreceptors in the eye of a mammal.
9. A recombinant adenovirus vector of claim 7, wherein the fiber is a
chimera composed of N-terminal sequences from adenovirus type 2 or type 5,
and a sufficient portion of an adenovirus serotype D fiber for selective binding to
photoreceptors in the eye of a mammal.
- 25 10. A method for targeted delivery of a gene product to the eye of a
mammal, comprising:
administering a recombinant adenovirus virus that comprises heterologous
DNA encoding the gene product or resulting in expression of the gene product,
wherein the recombinant virus comprises a fiber protein that specifically or
30 selectively binds to receptors that are expressed on cells in the eye.

-99-

11. The method of claim 10, wherein the cells are photoreceptors.
12. The method of claim 10, wherein administration is effected by intraocular delivery.
13. The method of claim 10, wherein administration is effected by a
5 method selected from subretinal injection, intravenous administration, periorbital administration, and intravitreal administration.
14. The method of claim 10, wherein the recombinant virus comprises a fiber protein from an adenovirus type D serotype.
15. The method of any of claims 10-14, wherein the fiber protein is an
10 adenovirus type 37.
16. The method of any of claims 10-14, wherein the fiber is a chimeric protein containing a sufficient portion of the N-terminus of an adenovirus type 2 or type 5 fiber protein for interaction with an adenovirus type 2 or type 5 penton, and a sufficient portion of an adenovirus serotype D knob portion of the
15 fiber for selective binding to photoreceptors in the eye of a mammal.
17. The method of any of claims 10-16, wherein the recombinant virus is an adenovirus type D serotype.
18. The method of any of claims 10-17, wherein the encapsulated nucleic acid comprises a photoreceptor-specific promoter operatively linked to a
20 nucleic acid comprising the therapeutic product.
19. The method of claim 18, wherein the therapeutic product is selected from the group consisting of a trophic factor, an anti-apoptotic factor, a gene encoding a rhodopsin protein, a wild-type Stargardt disease gene (STDG1), an anti-cancer agent and a protein that regulates expression of a photoreceptor-
25 specific gene product.
20. The method of any of claims 10-19, wherein delivery is effected for treatment of an ocular disease.
21. The method of claim 20, wherein the disorder is a retinal degenerative disease.
- 30 22. The method of claim 20, wherein the disease is retinitis pigmentosa, Stargardt's disease, diabetic retinopathies, retinal vascularization, or retinoblastoma.

-100-

23. The method of any of claims 10-22, wherein the mammal is a human.
24. The method of any of claims 10-22, wherein the viral nucleic acid comprises:
- 5 an adenovirus inverted terminal repeat (ITR) sequences; and
 an adenovirus packaging signal operatively linked thereto.
25. The method of claim 24, wherein the ITRs and packaging signal are derived from an adenovirus serotype B or C.
26. The method of claim 24, wherein the ITRs and packaging signal
10 are derived from an adenovirus type 2 or 5.
27. The method of claim 24, wherein the viral nucleic acid further comprises a photoreceptor-specific promoter.
28. A method of targeted gene therapy, comprising:
 administering a recombinant viral vector that comprises an adenovirus
15 type 37 fiber protein or portion thereof, whereby the vector selectively transduces photoreceptors and delivers a gene product encoded by the recombinant viral vector; wherein the portion is sufficient for selective binding to photoreceptors.
29. The method of claim 28, wherein the vector is administered into
20 the eye.
30. The method of claim 28, wherein the vector is administered to the vitreous cavity of the eye.
31. The method of claim 28, wherein administration is effected by subretinal injection, intravenous administration, periorbital administration or
25 intravitreal administration.
32. The method of any of claims 10-31, wherein at least about 10^7 plaque forming units of virus are administered.
33. The method of any of claims 10-31, wherein about 1 plaque forming unit to about 10^{14} plaque forming units of virus are administered.

SEQUENCE LISTING

<110> VON SEGGERN, DANIEL
NEMEROW, GLEN R.
FRIEDLANDER, MARTIN

<120> VECTORS FOR OCULAR TRANSDUCTION AND USE THEREFOR FOR GENETIC THERAPY

<130> 756.1PCT/NOV0205P

<140>

<141> 2001-05-01

<150> 09/562,934

<151> 2000-05-01

<160> 51

<170> PatentIn Ver. 2.1

<210> 1

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 1

cggtacacag aattcaggag acacaactcc

30

<210> 2

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 2

gcctggatcc gggaagttac gtaacgtggg aaaac

35

<210> 3

<211> 12

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: linker

<400> 3

cgcggtatccg cg

12

<210> 4

<211> 8710

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid

<400> 4
cacctaaatt gtaagcggtta atatttttgtt aaaattcgcg ttaaattttt gttaaattcag 60
ctcatttttt aaccaatagg ccgaaatcgg caaaatccct tataaatcaa aagaatagac 120
cgagataggg ttgagtggtt ttccagtttg gaacaagagt ccactattaa agaacgtgga 180
ctccaacgtc aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc 240
accctaataca agtttttttg ggtcgaggtg ccgtaaagca ctaaattcgg accctaaagg 300
gagccccga ttttagagctt gacggggaaa gccggcgaa gtggcgagaa aggaaggaa 360
gaaagcgaag ggagcgggag ctaggcgct ggcaagtgt gcggtcacgc tgcgcgtaac 420
caccacaccc gccgcgctta atgcgcgct acagggcgcg tcccattcgc cattcaggct 480
gcgcaactgt tgggaagggc gatcgggtcg ggcctcttcg ctattacgcc agtggcgaa 540
agggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg 600
ttgtaaaacg acggccagtg aattgtaata cgactcata tagggcgaat tgggtaccgg 660
gccccccctc gaggtcgacg gtatcgataa gcttgatata gaattcagga gacacaactc 720
caagtgcata ctctatgtca ttttcatggg actgggtctgg ccacaactac attaatgaa 780
tatttgccac atcctcttac actttttcat acattgccc agaataaaga atcgtttgg 840
ttatgtttca acgtgtttat ttttcaattg cagaaaattt caagtcattt ttcattcagt 900
agtatagccc caccaccaca tagcttatac agatcacctg accttaataa aactcacaga 960
accctagtat tcaacctgcc acctccctcc caacacacag agtacacagt cttttctccc 1020
cggctggcct taaaaagcat catatcatgg gtaacagaca tattcttagg tgttatattc 1080
cacacggttt cctgtcgagc caaacgctca tcagtgatat taataaactc cccggcgagc 1140
tcacttaagt tcatgtcgct gtccagctgc tgagccacag gctgctgtcc aacttgcgg 1200
tgcttaacgg gcggcgaaagg agaagtcac gcctacatgg gggtagagtc actgctgcc 1260
atcaggatag ggcgggtggt ctgcagcagc tctcagcga tgattcgac cgcccgcagc 1320
gtcctgcagg aatacaacat ggcagtggtc tctcagcga tgattcgac cgcccgcagc 1380
ataaggcgcc ttgtcctccg ggcacagcag cgcacctga tctcacttaa atcagcacag 1440
taactgcagc acagcaccac aatattgttc aaaatcccac agtgcaaggc gctgtatcca 1500
aagctcatgg cggggaccac agaaccacag tggccatcat accacaagcg caggtagatt 1560
aagtggcgac cctcataaa cacgctggac ataaacatta cctcttttgg catgtttaa 1620
ttcaccacct cccggtacca tataaacctc tgattaaaca tggcgccatc caccaccatc 1680
ctaaaccagc tggccaaaaac ctgcccgcgg gctatacact gcagggaacc gggactggaa 1740
caatgacagt ggagagccca ggactcgtaa ccatggatca tcatgctcgt catgatata 1800
atgttggcac aacacaggca cacgtgcata cacttcctca ggattacaag ctccctccgc 1860
gttagaacca tatcccaggg aacaacccat tctgtaata gcgtaaatcc cacactgcag 1920
ggaagacctc gcacgtaact cacgttgtgc tctgtctcaa aaggaggtag acgatcccta 2040
ggatgatcct ccagtatggt agcgcgggtt tctgtctcaa aaggaggtag acgatcccta 2040
ctgtacggag tgcccgagga caaccgagat cgtgttggtc gtagtgctat gccaaatgga 2100
acgcccgaag tagtcatatt tctgaagca aaaccagtg cgggcgtgac aaacagatct 2160
gcgtctccgg tctgcgcgt tagatcgctc ttgttagtag ttgtagtata tccactctct 2220
caaagcatcc aggcgcccc tggcttcggg ttctatgtaa actccttcat gcgcccgtgc 2280
cctgataaca tccaccaccg cagaataagc cacaccagc caacctacac attcgttctg 2340
cgagtccac acgggaggag cgggaagagc ttggaagaacc atgtttttt ttttattcca 2400
aaagattatc caaaacctca aaatgaagat ctattaagt aacgcgctcc cctccgggtg 2460
cgtgggtcaaa ctctacagcc aaagaacaga taatggcatt tgtaagatgt tgcacaatgg 2520
cttccaaaag gcaaacggcc ctcacgtcca agtggacgta aaggctaaac ccttcagggt 2580
gaatctcctc tataaacatt ccagcacctt caaccatgcc caaataattc tcatctcgcc 2640
accttctcaa tatatctcta agcaaatccc gaatatatt tccggccatt gtaaaaaatc 2700
gctccagagc gccctccacc ttcagcctca agcagcgaat catgattgca aaaattcagg 2760
ttcctcacag acctgtataa gattcaaaag cggaaacatta acaaaaatac cgcgatccc 2820
taggtccctt cgcagggcca gctgaacata atcgtgcagg tctgcacgga ccagcgcgcc 2880
cacttccccg ccaggaacct tgacaaaaga acccacactg attatgacac gcatactcgg 2940
agctatgcta accagcgtag ccccgatgta agctttgttg catggcgggc gatataaaat 3000
gcaaggtgct gctcaaaaaa tcaggcaaa cctcgcgcaa aaaagaaagc acatcgtagt 3060
catgctcatg cagataaagg caggtaagct ccggaaccac cacagaaaaa gacaccattt 3120
ttctctcaaa catgtctgag ggtttctgca taacacaaaa ataaaaatac aaaaaaacat 3180
ttaaacatta gaagcctgtc ttacaacagg aaaaacaacc cttataagca taagacggac 3240
tacggccatg ccggcgtgac cgtaaaaaaa ctggtcaccg tgattaaaaa gcaccaccga 3300
cagtcctcag gctcatgtcc gagtcataat gtaagactcg gtaaacacat caggttgatt 3360
catcggtcag tgctaaaaag cgaccgaaat agcccggggg aatacatacc cgcaggcgta 3420
gagacaacat tacagcccc ataggaggta taacaaaatt aataggagag aaaaacacat 3480
aaacacctga aaaacctcc tgcctaggca aaatagcacc ctcccgtctc agaacaacat 3540

acagcgcttc acagcgggcag cctaacagtc agccttacca gtaaaaaaga aaacctatta 3600
aaaaaacacc actcgacacg actaaaaaat aatcagtcac agtgtaaaaa agggccaagt 3660
gcagagcgag tatatatagg actaaacacgg gacgtaacgg ttaaaagtcca caaaaaacac 3720
ccagaaaaacc gcacgcgaac ctacgcccag aaacgaaagc caaaaaaccc acaacttcct 3780
caaactcgta cttccggtttt ccacagttac gtaacttccc ggatccgagg cattcacagt 3840
tctccgcaag aattgattgg ctccaattct tggagtgggt aatccgttag cgaggcgccg 3900
ccggcttcca ttcaggtcga ggtggccggg ctccatgcac cgcgacgcaa cgcggggagg 3960
cagacaaggt atagggcggc gcctacaatc catgccaacc cgttccatgt gctcgccgag 4020
gcggcataaa tcgcccgtgac gatcagcggt ccagtgtatg aagttaggct ggtaagagcc 4080
gcgagcgatc cttgaagctg tccctgatgg tcgtcatcta cctgcctgga cagcatggcc 4140
tgcaacgcgg gcatcccgat gccgcgggaa gcgagaagaa tcataatggg gaaggccatc 4200
cagcctcgcg tcgcgaacgc cagcaagacg tagcccagcg cgtcggccgc catgccctgc 4260
ttcatccccg tggcccggtt ctccgctttg ctggcggtgt ccccgaaga aatatatttg 4320
catgtcctta gttctatgat gacacaaacc ccgcccagcg ccttcgcatat taagggtgacg 4440
aacacgcaga tgcagtcggg cgggcgcggg cccaggtcca cttcgcatat caacagcgtg 4500
cgtgtggcct cgaacaccga gcgaccctgc agcgaccgcg caccgcgacg tctgtcgaga 4560
ccgcagatcc cgggcaatga gatatgaaaa agcctgaact gcagctctcg gagggcgag 4620
agtttctgat cgaaaagtgc gacagcgtct ccgacctgat tgtcctgcgg gtaaatagct 4680
aatctcgtgc tttcagcttc gatgtaggag ggcgtggata ctttgcacg gcccgctcc 4740
gcgcccgatg tttctacaaa gatcgttatg tttatcggca ctttgcacg gctgtctccc 4800
cgattccgga agtgcttgac attggggaaat tcagcgagag cctgaactat gctgttctgc 4860
gccgtgcaca ggggtgtcac ttgcaagacc ctgcggccga tcttagccag acgagcggtg 4920
agccggtcgc gggaggccatg gatgcgatcg ggaatcggtc aatacactac atggcgtgat 4980
tcggcccatt cggaccgcaa ggaatcggtc aaactgtgat ggactgcccc gaagtccggc 5040
cgattgtgta tccccatgtg tatcactggc tttggggcga ggactgcccc gcataacag 5100
ccgtcgcgca ggctctcgat ggctccaaca atgtcctgac ggacaatggc gccaacatct 5160
acctcgtgca cgcggatttc gcgatgttcg gggattccca ggcactactc gagcggaggc 5280
cggctcattga ctggagcgag gcttgtatgg ggcagcagag gctccgcatt ggtcctgacc 5340
tcttctggag gccgtgggtg ccgcggctcc gggcgatat tgcgatgac agcttgggag 5400
atccggagct gagcttgggt gacggcaatt ctgtcggcg tacacaaatc gcccgagaa 5460
aactctatca cgtccgatcc ggagccggga aagtagtcgc cgatagtgga aaccgacgcc 5520
gcgagcgaat ctggacgat tccgagggca aaggaatagg ggagatgggg gaggttaact gaaacacgga 5580
ccagcactcg accggaagga acccgcgcta tgacggcaat aaaaagacag aataaaacgc 5640
aggagacaat ggtcgtttgt ttggggccaa tacgcccgcg tttcttccct ttccccaccc 5700
acgggtgttg cgagaccoca gttcgggtga agggccaggg ctgcgagcca acgtcggggc ggcaggccct 5760
gatacccccac gttcgggtga agggccaggg cggggctccc catggggaat ggtttatgg 5820
caccccccaa ctggccccgt ggggttaggga ggggtctggt ccacgactgg actgagcaga 5940
gccatagcca tattttttg ggcgttgcgt tggaccgcat gtactggcgc gacacgaaca 6000
tcgtgggggt gtttttggat ggctgggca aacacccccg acccccacac ccgctgacgc 6060
cagacccatg gtggctgcca cgtcgaccgg gcccgcacac ccgccaacac ccgtctccgg 6120
ccgggcgtct cgtcgaccgg gcttgtctgc tcccggcatc cgcttacaga caagctgtga cgcgcgaggc 6180
gccttgacgg gcttgtctgc tttcaccgtc atcaccgaaa cgcgcgaggc agccggatca 6240
gagctgcatg tgtcagaggt tttcaccgtc tacttgcctt aaaaaacctc cccactccc 6300
taatcagcca taccacattt gtagaggttt tgttgttgg aacttgttta ttgcagctta 6360
cctgaacctg aaacataaaa tgaatgcaat aaatttcaca aataaagcat ttttttcaact 6420
taatggttac aaataaagca atagcatcac caatgtatct tatcatgtct ggatccacta 6480
gcattctagt tgtggtttgt ccaaactcat ctagctttt gttcccttta gtgagggtta 6540
gttctagagc ggccgcccac atggtcatag ctgtttcctg tgtgaaattg ttatccgctc 6600
atttcgagct tggcgtaatc agccggaagc ataaaagtgt aagcctgggg ctttccagtc 6660
acaattccac acaacatacg atggctgacg cgcgcgggga ctttccagtc gggaaaactg 6720
gtgagctaac tcacattaat tgcgttgcgc tcaactgccc cgttccagtc ggcgtattggg 6780
tcgtgccagc tgcattaatg aatcggccaa cgtcgctcgg cgttccagtc ggcggcagcg 6840
cgctcttccg cttcctcgct cactgactcg tttatccacag aatcagggga taacgcagga 6900
gtatcagctc actcaaaggc ggtaatacgg gccaggaacc gtaaaaaagg ccggttgctg 6960
aagaacatgt gagcaaaaagg ccccccgtgac taccaggcgt ttccccctgg aagctccctc 7020
cggtttttcc ataggctccg actataaaga acccgatacc tgteccgctt tctcccttcg 7080
aggtggcgaa ctgttccgac cctgcgcgtt tctaggtatc cagttccgtg gtaggtcgt 7140
gtgcgctctc cgctttctca tagctcacgc cccgttcagc cgcacttatcc 7200
ggaagcgtgg cgtccaagc ctaacagtc aatcagtcac agtgtaaaaa agggccaagt 7260

```

ggtaactatc gtcttgagtc caaccccggt agacacgact tatcgccact ggcagcagcc 7320
actggtaaca ggattagcag agcgaggat gtaggcggtg ctacagagtt cttgaagtgg 7380
tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 7440
gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggttagc 7500
ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 7560
cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg ttaagggatt 7620
ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 7680
tttaaatcaa tctaaagtat atatgagtaa acttggcttg acagttacca atgcttaatc 7740
agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagtgc ctgactcccc 7800
gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata 7860
ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 7920
gccgagcgca gaagtgggcc tgcaacttta tccgcctcca tccagtctat taattggtgc 7980
cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct 8040
acaggcatcg tgggtgcacg ctctcggtt ggtaggctt cattcagctc cggttcccaa 8100
cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcgggttag ctcttcggt 8160
cctccgatcg ttgtcagaag taagtgggcc gcagtggtat cactcatggt tatggcagca 8220
ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tgggtgagta 8280
tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgctca 8340
atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt 8400
tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaacct 8460
actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgttc tgggtgagca 8520
aaaacaggaa ggcaaaaatgc cgcaaaaaag ggaataaggc cgacacggaa atgttgaata 8580
ctcatactct tcctttttca atattattga agcatttatc aggggttattg tctcatgagc 8640
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 8700
cgaaaagtgc

```

<210> 5
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 5
 atgggatcca agatgaagcg cgcaagaccg 30

<210> 6
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 6
 cataacgcgg ccgcttcttt attcttgggc 30

<210> 7
 <211> 7148
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: plasmid

<400> 7
 gacggatcgg gagatctccc gatccctat ggtcgactct cagtacaatc tgctctgatg 60
 ccgcatagtt aagccagat ctgctccctg cttgtgtggt ggaggtcgct gagtagtgcg 120
 cgagcaaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180

ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatagc cgttgacatt 240
gattattgac tagttattaa tagtaataca ttacgggggc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc 360
cccggccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggg aaactgccc cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagta 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggag 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttgggtacc 900
gagctcggat ccaagatgaa gcgcgcaaga ccgtctgaag ataccttcaa ccccggtgat 960
ccatagaca cggaaaccgg tcctccaact gtgccttttc ttactcctcc ctttgtatcc 1020
cccaatgggt ttcaagagag tccccctggg gtactctctt tgcgcctatc cgaacctcta 1080
gttacctcca atggcatgct tggcgtcaaa atgggcaacg gcctctctct ggacgaggcc 1140
ggcaacctta cctcccaaaa tgtaaccact gtgagccac ctctcaaaaa aaccaagtca 1200
aacataaacc tggaaatatc tgcacccctc acagttacct cagaagccct aactgtggct 1260
gccgcccac ctctaattgg tagcattgac accactacca ccctcacagt gtcagaagga 1320
accgtgcacg actccaaact agggccccct accaccaccg atagcagtag ccttactatc 1380
aagctagccc tgcaaacatc agggccccct ggtagcttgg gcattgactt gaaagagccc 1440
actgcctcac cccctctaac tactgccact aagtacgggg ctcctttgca tgtaacagac 1500
atctatacac aaaatggaaa actaggacta ccaggtgtga ctattaataa tacttccctg 1560
gacctaaaca ctttgaccgt agcaactggg gattcacaag gcaatatgca acttaagtga 1620
caaactaaag ttactggagc cttgggtttt agagcctta tacttgatgt tagttatccg 1680
gcaggaggac taaggattga ttctcaaaac agagcctta gacctctttt tataaactca 1740
tttgatgctc aaaaccaact ctacaacaaa ggcctttact tgtttacagc ttcaaaacaat 1800
gcccacaact tggatattaa cctaagcact gccaaggggt tgatgtttga cgctacagcc 1860
tccaaaaagc ttgaggttaa cttgggtcac ctaaatgcacc aaacacaaaat 1920
atagccatta atgcaggaga tgggcttgaa ttgggttgc caaacaaggc tatgggtcct 1980
ccctcaaaa ctggccttag ttttgacagc acaggtgcca ttacagtagg aaacaaaaat 2040
aaactaggaa aatgataagc taacttttg gaccacacca gctccatctc ctaactgtag acttgctaca 2100
gagaaagatg gtttcagttt tggctgttaa aggcagtttg gctccaatat ctggaacagt tcaaaagtgt 2160
catcttatta taagatttga cgaataatgga gtgctactaa acaattcctt cctggaccga 2220
gaatatttga actttagaaa tggagatctt actgaaggca cagcctatac aaacgtgtgt 2280
ggattttatg ctaacctatc agcttatcca aaatctcac gtaaaactgc caaaagtaac 2340
attgtcagtc aagtttactt aaacggagac aaaactaaac ctgtaacact aaccattaca 2400
ctaaacggta cacaggaaac aggagacaca actccaagt gtaaaactgc aacatttca 2460
tgggactggg ctggccacaa ctacattaat gaaatatttg catactctat gtcattttca 2520
tcatacattg cccaagaata aagaagcggc gctcgagca ccacatctc ttacactttt 2580
ctatagtgtc acctaaatgc tagagctcgc tgatcagcct tgcattctaga gggccctatt 2640
cagccatctg ttgtttgccc ctccccctg ccttccctga cactgtgccc ttctagtgtc 2700
actgtccttt cctaataaaa tgaggaaatt gcagacagc ccttccctga cctggaagg tgccactccc 2760
attctggggg gtgggggtggg gcaggacagc aagggggagg gtctgagtag gtgtcattct 2820
catgctgggg atgcgggtggg ctctatggct aagggggagg attgggaaga caatagcagg 2880
aggggggtat cccacgcgcc cgttagcgcc gcattaaagc cggcgggtgt ctgggttacg 2940
cgcagcgtga ccgctacact tgccagcgcc ctagecgccc ctcctttcgc tttcttccct 3000
tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg catcccttta 3060
gggttccgat ttagtgcttt acggcacctc ctgatagacg gtttttgcgc ctttgacgtt gggtgatgg 3120
tcacgtagtg ggccatcgcc gttccaaact ggaacaacac tcaaccctat ctcgggtctat 3180
ttctttaata gtggactctt tttggggatt tggcctatt ggttaaaaaa tgagctgatt 3240
tcttttgatt tataagggat ttaattctgt ggaattgtgt tcaagttagg ggttaaaaaa 3300
taacaaaaat ttaacgcgaa cagaagtatg caaagcatgc atctcaatta gtcagcaacc 3360
ccccaggctc agtccccagg ctccccagca ggcagaagta tgcaaaagcat gcatctcaat 3420
aggtgtggaa ccatagtccc gcccttaact ccgcccaccc cggcccctaac tccgcccagg 3480
tagtcagcaa ctccgccccca tggctgacta atttttttta tttttggagg cctaggcttt 3540
tccgcccatt ctccgccccca ccagaagtag tgaggaggct tttttggagg gacaggatga 3600
gcctctgcct ctgagctatt ttgtatatcc attttcggat ctgatcaaga cgcttgggtg 3660
tgcaaaaagc tcccggggagc acaagatgga ttgcacgcag gttctccggc cgttgggtg 3720
ggatcgcttc gcatgattga

```

gagaggctat tccggctatga ctggggcacia cagacaatcg gctgctctga tgccgcccgtg 3960
ttccggctgt cagcgcaggg gcgcccgggt ctttttgtca agaccgacct gtccgggtgcc 4020
ctgaatgaac tgcaggacga ggcagcgcgg ctatcgtggc tggccacgac gggcgcttcc 4080
tgccgacgtg tgctcgacgt tgtcactgaa gcgggaaggg actggctgct attgggcgaa 4140
gtgccggggc aggatctcct gtcatctcac cttgtcctg cggagaaagt atccatcatg 4200
gctgatgcaa tgcggcgggt gcatacgtt gatccggcta cctgccatt cgaccaccaa 4260
gcgaaacatc gcatcgagcg agcacgtact cggatggaag ccggtcttgt cgatcaggat 4320
gatctggacg aagagcatca ggggctcgcg ccagccgaac tgttcgccag gctcaaggcg 4380
cgatgccccg acggcgagga tctcgtcgtg acccatggcg atgcctgctt gccgaatatc 4440
atgggtgaaa atggccgctt ttctggatc atcgactgtg gccggctggg tgtggcggac 4500
cgctatcagg acatagcgtt ggctaccctg gatattgctg aagagcttgg cggcgaatgg 4560
gctgaccgct tcctcgtgct ttacggtatc ccgctcccg attcgcagcg catcgcttcc 4620
tatcgcttcc ttgacgagtt cttctgagcg ggactctggg gttcgaaatg accgaccaag 4680
cgacgcccac cctgccatca cgagatttct attccaccgc cgccttctat gaaagggttg 4740
gcttcggaat cgttttccgg gacgcggct ggatgatcct taatggttac aaataaagca 4800
tggagtctct cgccaccccc aacttgttta ttgcagctta taatggttac aaataaagca 4860
atagcatcac aaatttcaca aataaagcat ttttttact gcattctagt tgtggtttgt 4920
ccaaactcat caatgtatct tatcatgtct gtataccgtc gacctctagc tagagcttgg 4980
cgtaatcatg gtcatactgt tttcctgtgt gaaattgtta tccgctcaca attccacaca 5040
acatacgagc cggaaagata aagtgtaaag cctgggggtgc ctaatgagtg agctaaactca 5100
cattaattgc gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc 5160
attaatgaat cggccaacgc gcggggagag gcggtttgct tattgggcgc tcttccgctt 5220
cctcgctcac tgactcgtg tccacagaat caggggataa cgcaggaaag aacatgtgag 5340
caaaaggcga gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 5400
ggctccgccc cctgacgag catcacaaaa atcgacgtc cccctggaag ctccctcgtg 5460
cgacaggact ataaagatac caggcgcttc cccctggaag cccttcggga agcgtggcgc 5520
ttccgaccct gccgcttacc ggatacctgt ccgcttttct cccttcggga agcgtggcgc 5580
tttctcaatg ctcacgctgt aggtatctca gttcgggtga ggtcgttcgc tccaagctgg 5640
gctgtgtgca gtaaccccc gttcagccc accgctgcgc cttatccggg aactatcgtc 5700
ttgagtcctc cccggtaaga cactgcttat cgcactggc agcagccact ggtaacagga 5760
ttagcagagc gaggtatgta ggcgggtgta cagagttctt gaagtgggtg cctaactacg 5820
gctacactag aaggacagta tttgggtatc gcgctctgct gaagccagtt accttcgga 5880
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg gggttttttg 5940
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 6000
ctacgggggtc tgacgctcag tggaaacgaa actcacgtta agggattttg gtcattgagat 6060
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 6120
aaagtatata tgagtaaaact tggctctgaca gttaccaatg cttaatcagt gaggcaccta 6180
tctcagcgat ctgtctatct cgttcatcca tagttgctg actccccgtc gtgtagataa 6240
ctacgatacg ggagggttta ccactctggc ccagtgtgc aatgataccg cgagaccac 6300
gctcacccgc tccagattta tcagcaataa accagccagc cggaaagggc gagcgagaa 6360
gtggctcctg aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag 6420
taagtagttc gccagttaat agtttgcgca acgttgttgc cattgtctca ggcacgtgg 6480
tgtcacgctc gtcgtttggg atggcttcat tcagctccgg ttcccaacga tcaaggcgag 6540
ttacatgata ccccatgttg tgcaaaaaag cggttagctc cttcggctct ccgatcgttg 6600
tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc 6660
ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtaacta accaagtcac 6720
tctgagaata gtgtatgcg cgaccgagtt gctcttgccc ggcgtcaata cgggataata 6780
ccgcgcccac tagcagaact ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa 6840
aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaaccactc cgtgcaccca 6900
actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaagg 6960
aaaaatgcgc aaaaaaggga ataaggcgga caggaaatg ttgaatactc atactcttcc 7020
tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg 7080
aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgcac 7140
ctgacgtc 7148

```

<210> 8

<211> 7469

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid

<400> 8

```
gacggatcgg gagatctccc gatccccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctcccctg cttgtgtgtt ggaggctcgt gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgtttttgc ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacgggggc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgccatt gagctcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgccc cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt cttccaccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acgggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc 900
gagctcggat ctgaattcga gctcgctgtt gggctcgcg ttaggacaa actottcgcg 960
gtctttccag tactcttggg tcggaaaccc gtcggcctcc gaacgggtact ccgccaccga 1020
gggacctgag cgagtcgcga tcgaccggat cggaaaacct ctcgagaaag gcgtctaacc 1080
agtcacagtc gcaaggtagg ctgagcaccg tggcggcgcg cagcgggttg cggctcggggt 1140
tggtttctggc ggaggtgctg ctgatgatgt aattaaagta ggcggtcttg agacggcgga 1200
tggtcgaggt gaggtgtggc aggtctgaga tccaagatga agcgcgcaag accgtctgaa 1260
gataccttca acccgtgta tccatagac acgaaaaccg gtccctcaac tgtgaccttt 1320
cttactctc cctttgtatc cccaatggg tttcaagaga gtccccctgg ggtactctct 1380
ttgcgcctat ccgaacctct agttacctcc aatggcatgc ttgcgctcaa aatgggcaac 1440
ggcctctctc tggacgagtc cggcaacctt acctcccaaa atgtaaccac tbtgagccca 1500
cctctcaaaa aaaccaagtc aaacataaac ctggaaatat ctgacacct caccagttacc 1560
tcagaagccc taactgtggc tgccgcccga cctctaattg tcgcgggcaa cacactcacc 1620
atgcaatcac agggcccgtt aaccgtgcac gactccaaac ttagcattgc cacccaagga 1680
cccctcacag tgtcagaagg aaagctagcc ctgcacaacat caggccccct caccaccacc 1740
gatagcagta ccttactat cactgcctca cccctctaa ctactgccac tggtagcttg 1800
ggcattgact tgaagagcc catttataca caaatggaa aactaggact aaagtacggg 1860
gctcctttgc atgtaacaga cgacctaaac actttgaccg tagcaactgg tccaggtgtg 1920
actattaata atacttctt gcaaaactaaa gttactggag ccttgggttt tgattcacia 1980
ggcaatatgc aacttaattg agcaggagga ctaaggattg attctcaaaa cagacgcctt 2040
atacttgatg ttagttatcc gtttgatgct caaaaccaac taaatctaag actaggacag 2100
ggcctctctt ttataaactc agcccacaac ttggatatta actacaacaa aggcctttac 2160
ttgtttacag cttcaaaaca ttccaaaaag cttgaggtta acctaaagcac tgccaagggg 2220
ttgatgtttg acgctacagc catagccatt aatgcaggag atggggttga atttggttca 2280
cctaattgcac caaacacaaa tcccctcaaa acaaaaattg gccatggcct agaatttgat 2340
tcaaaacagg ctatggttcc taaactagga actggcctta gttttgacag cacaggtgcc 2400
attacagtag gaaacaaaaa taatgataag ctaactttgt ggaccacacc agctccatct 2460
cctaactgta gactaaatgc agagaaagat gctaaactca ctttggctct aacaaaatgt 2520
ggcagtcaaa tacttgctac agtttcagtt ttggctgtta aaggcagttt ggctccaata 2580
tctggaacag ttcaaagtgc tcatcttatt ataagatttg acgaaaatgg agtgctacta 2640
aacaattcct tcctggacc cagaatattg aactttagaa atggagatct tactgaaggc 2700
acagcctata caaacgctgt tggatttatg cctaacctat cagcttatcc aaaatctcac 2760
ggtaaaactg ccaaaagtaa cattgtcagt caagtttact taaacggaga caaaactaaa 2820
cctgtaaacac taaccattac actaaacggt acacaggaaa caggagacac aactccaagt 2880
gcatactcta tgtcattttc atgggactgg tctggccaca actacattaa tgaaatattt 2940
gccacatcct cttacacttt ttcatacatt gcccaagaat aaagaagcgg ccgctcgagc 3000
atgcatctag agggccctat tctatagtgt cacctaaatg ctagagctcg ctgatcagcc 3060
tcgactgtgc cttctagttg ccagccatct gttgtttgcc cctccccctg gccttccttg 3120
accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcacgcgat 3180
tgtctgagta ggtgtcattc tattctgggg ggtgggggtg ggcaggacag caagggggag 3240
gattgggaag acaatagcag gcatgctggg gatcggttgg gctctatggc ttctgaggcg 3300
gaaagaacca gctgggggtc taggggggtat cccacgcgc cctgtagcgg cgcattaagc 3360
gcggcggggtg tgggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc 3420
gtcctcttgc ctttcttccc ttcctttctc gccacgttcg ccggctttcc ccgtcaagct 3480
```

ctaaatcggg	gcatcccttt	agggttccga	tttagtgctt	tacggcacct	cgaccccaaa	3540
aaacttgatt	aggggtgatg	ttcacgtagt	gggcatcg	cctgatagac	gggtttttcg	3600
cctttgacgt	tggagtcac	gttctttaat	agtggactct	tggtccaaac	tggacaaca	3660
ctcaacccta	tctcggtcta	ttcttttgat	ttataagga	ttttggggat	ttcgccctat	3720
tggttaaaaa	atgagctgat	ttacaaaaa	tttaacgcga	attaattctg	tggaaatgtg	3780
gtcagttagg	gtgtggaaa	tccccaggct	ccccaggcag	gcagaagtat	gcaaagcatg	3840
catctcaatt	agtacgaac	caggtgtgga	aagtcaccag	gctccccagc	aggcagaagt	3900
atgcaaaagc	tgcatctcaa	ttagtcagca	accatagtc	cgccccctaac	tccgcccctc	3960
ccgcccctaa	ctccgcccag	ttccgcccct	tctccgcccc	atggctgact	aatttttttt	4020
atttatgcag	aggccgaggc	cgcctctgcc	tctgagctat	tccagaagta	gtgaggaggc	4080
ttttttggag	gcctaggctt	ttgcaaaaag	ctcccgagg	cttgatatc	catttttcgga	4140
tctgatcaag	agacagatg	aggatcggtt	cgcattgatg	aacaagatgg	attgcacgca	4200
gggtctccgg	ccgcttgggt	ggagaggcta	ttcggtatg	actgggcaca	acagacaatc	4260
ggctgctctg	atgcccgctg	gttccggctg	tcagcgcagg	ggcgcccggt	tctttttgtc	4320
aagaccgacc	tgcccggtgc	cctgaatgaa	ctgcaggacg	aggcagcgcg	gctatcgtgg	4380
ctggccacga	cgggcgttcc	ttgctgcagc	gtgctgcagc	ttgtcactga	agcgggaagg	4440
gactggctgc	tattgggcga	agtgcggggg	caggatctcc	tgcatctca	ccttgctcct	4500
gccgagaaa	tatccatcat	ggctgatgca	atgcggcgcc	tgcatagct	tgatccggct	4560
acctgcccct	tcgaccacca	agcgaaacat	cgcctcgagc	gagcacgtac	tcggatggaa	4620
gccggtcttg	tcgatcagga	tgatctggac	gaagagcatc	aggggctcgc	gccagccgaa	4680
ctgttcgcca	ggctcaaggc	gcgcatgccc	gacggcgagg	atctcgtcgt	gacccatgge	4740
gatgcctgct	tgccgaatat	catgggtgaa	aatggccgct	tttctggatt	catcgactgt	4800
ggccggctgg	gtgtggcgga	ccgctatcag	gacatagcgt	tggtaccctg	tgatattgct	4860
gaagagcttg	gcggcgcaatg	ggctgaccgc	ttcctcgtgc	tttacgggtat	cgccgctccc	4920
gattcgcagc	gcattgcctt	ctatgcctt	cttgacgagt	tcttctgagc	gggactctgg	4980
ggttcgaaat	gaccgaccaa	gcgagcccca	acctgccatc	acgagatttc	gattccaccg	5040
ccgccttcta	tgaaagggtg	ggcttcggaa	tcggtttccg	ggacgcgggc	tggtatgatcc	5100
tccagcgcg	ggatctcatg	ctggagttct	tcgcccaccc	caacttggtt	attgcagctt	5160
ataatggtta	caaataaagc	aatagcatca	caaatttcac	aaataaagca	tttttttcac	5220
tgcatctag	ttgtgggttg	tcctaaactca	tcaatgtatc	ttatcatgtc	tgtataccgt	5280
cgacctctag	ctagagcttg	gcgtaatcat	ggcatagct	gtttcctgtg	tgaaattggt	5340
atccgctcac	aattccacac	aacatacgag	ccggaagcat	aaagtgtaaa	gcctgggggtg	5400
cctaagtagt	gagctaaact	acattaatgg	cgttgcgctc	actgcccgct	ttccagtcgg	5460
gaaacctgtc	gtgcccagctg	cattaatgaa	tcggccaacg	cgcggggaga	ggcggtttgc	5520
gtattgggctg	ctcttccgct	tctctgctca	ctgactcgct	gcgctcggtc	gttcgggctgc	5580
ggcgagcggt	atcagctcac	tcaaaggcgg	taatacgggt	atccacagaa	tcaggggata	5640
acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	5700
cggtgctggc	gtttttccat	aggctccgcc	cccctgacga	gcatcacaaa	aatcgacgct	5760
caagtgcag	gtggcgaaac	ccgacaggac	tataaagata	ccaggcggtt	ccccctggaa	5820
gtccctcgt	gcgctctcct	gttccgaccc	tgccgcttac	cggataacctg	tccgcttttc	5880
tcccttcggg	aagcgtggcg	ctttctcaat	gctcacgctg	taggtatctc	agttcggtgt	5940
aggctcgttcg	ctccaagctg	ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgcg	6000
ccttatccgg	taactatcgt	cttgagtcca	acccggtaag	acacgactta	tcgccactgg	6060
cagcagccac	tggtaacagg	attagcagag	cgaggtatgt	aggcggtgct	acagagtctc	6120
tgaagtgggtg	gcctaactac	ggctacacta	gaaggacagt	atgttggtatc	tgcgctctgc	6180
tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaa	caaaccaccg	6240
ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	gcgcagaaaa	aaaggatctc	6300
aagaagatcc	tttgatcttt	tctacggggg	ctgacgctca	gtggaacgaa	aactcacggt	6360
aagggatttt	ggctatgaga	ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	6420
aatgaagttt	taaatcaatc	taaagtatat	atgagtaaac	ttggtctgac	agttaccaat	6480
gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	tcggtcatcc	atagttgcct	6540
gactccccgt	cgtgtagata	actacgatac	gggagggctt	accatctggc	cccagtgtctg	6600
caatgatacc	gcgagaccca	cgctcaccgg	ctccagattt	atcagcaata	aaccagccag	6660
ccggaagggc	cgagcgcaga	agtggctcctg	caactttatc	cgctccatc	cagtcctatta	6720
attgttgccg	ggaagctaga	gtaagttagt	cgccagttaa	tagtttgccg	aacgttggtg	6780
ccattgctac	aggcatcggtg	gtgtcacgct	cgtcggttgg	tatggcttca	ttcagctccg	6840
gttcccaacg	atcaaggcga	gttacatgat	ccccatggt	gtgcaaaaaa	gcgggttagct	6900
ccttcgggtcc	tccgatcggt	gtcagaagta	agttggccgc	agttgtatca	ctcatgggta	6960
tgccagcaact	gcataattct	cttactgtca	tgccatccgt	aagatgcttt	tctgtgactg	7020
gtgagtactc	aaccaagtca	ttctgagaat	agtgatgcg	gcgaccgagt	tgctcttgcc	7080
cgccgtcaat	acgggataat	accgcgccac	atagcagaac	tttaaaagtg	ctcatcattg	7140
gaaaacgttc	ttcggggcga	aaactctcaa	ggatcttacc	gctggttgaga	tccagttcga	7200


```

tgtaaccacac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcggtttctg 7260
gggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat 7320
ggtgaatact catactcttc ctttttcaat attattgaag catttatcag gggtattgtc 7380
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataagg gttccgcgca 7440
catttccccg aaaagtgcc cctgacgtc 7469

```

<210> 9
 <211> 28
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 9
 tgcttaagcg gccgcgaagg agaagtcc 28

<210> 10
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 10
 ccgagctagc gactgaaaat gag 23

<210> 11
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 11
 cctctcgaga gacagcaaga cac 23

<210> 12
 <211> 11152
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: plasmid

<400> 12
 aagcttgggc agaaatgggt gaactccgga gagtgtccta cacctagggg agaagcagcc 60
 aaggggttgt ttcccaccaa ggacgacccg tctgcgcaca aacggatgag cccatcagac 120
 aaagacatat tcattctctg ctgcaaaactt ggcataagctc tgctttgcct ggggctattg 180
 ggggaagttg cggttcgtgc tcgcagggct ctcacccttg actcttttaa tagctcttct 240
 gtgcaagatt acaatctaaa caattcggag aactcgacct tcctcctgag gcaaggacca 300
 cagccaactt cctcttaca ggcgcacgca tttgtcctt cagaaataga aataagaatg 360
 cttgctaataa attatatttt taccaataag accaatccaa taggtagatt attagttact 420
 atgttaagaa atgaatcatt atcttttagt actattttta ctcaaattca gaagttagaa 480
 atgggaatag aaaatagaaa gagacgctca acctcaattg aagaacagggt gcaaggacta 540
 ttgaccacag gcctagaagt aaaaaaggga aaaaagagtg tttttgtcaa aataggagac 600
 aggtgggtggc aaccagggac ttatagggga cttacatct acagaccaac agatgcccc 660
 ttaccatata caggaagata tgacttaaat tgggataggt ggggtacagt caatggctat 720

aaagtgttat	atagatccct	cccttttctg	gaaagactcg	ccagagctag	acctccttgg	780
tgtatgttgt	ctcaagaaga	aaaagacgac	atgaaacaac	aggtacatga	ttatatattat	840
ctaggaacag	gaatgcactt	ttggggaaag	attttccata	ccaaggaggg	gacagtggct	900
ggactaatag	aacattatct	tgcaaaaact	catggcatga	gttattatga	atagccttta	960
ttggcccaac	cttgcgggtc	ccagggctta	agtaagtttt	tggttacaaa	ctgttcttaa	1020
aacgaggatg	tgagacaagt	ggtttcctga	cttgggttgg	tatcaaaggt	tctgatctga	1080
gctctgagtg	ttctattttc	ctatgttctt	ttggaaattt	tccaaatctt	atgtaaatgc	1140
ttatgtaaac	caagatataa	aagagtgtcg	attttttgag	taaaacttgc	acagtcctaa	1200
cattcacctc	ttgtgtgttt	gtgtctgttc	gccatcccgt	ctccgctcgt	cacttatcct	1260
tcaactttcca	gaggggtccc	ccgcagaccc	cggcgaccct	caggtcggcc	gactgcgcca	1320
gctggcgccc	gaacagggac	cctcggataa	gtgacccttg	tctctatttc	tactatttgg	1380
tgtttgtctt	gtattgtctc	tttcttgtct	ggctatcatc	acaagagcgg	aacggactca	1440
ccatagggac	caagctagcg	actgaaaatg	agacatatta	tctgccacgg	aggtgttatt	1500
accgaagaaa	tgcccgccag	tcttttgagc	cagctgatcg	aagaggtaet	ggctgataat	1560
cttccacctc	ctagccattt	tgaaccacct	acccttcacg	aactgtatga	tttagacgtg	1620
acggcccccg	aagatcccaa	cgaggaggcg	gtttcgcaga	tttttcccga	ctctgtaatg	1680
ttggcggtgc	aggaagggat	tgacttactc	acttttccgc	cggcgcccgg	ttctccggag	1740
ccgcctcacc	tttcccgcca	gcccagagcag	ccggagcaga	gagccttggg	tccgggttct	1800
atgccaaacc	ttgtaccgga	ggtgatcgat	cttacctgcc	acgaggctgg	ctttccaccc	1860
agtgcagacg	aggatgaaga	gggtgaggag	tttgtgttag	attatgtgga	gcaccccggg	1920
cacggttgca	ggtcttgtca	ttatcacccg	aggaatacgg	gggaccacga	tattatgtgt	1980
tgcctttgct	atatgaggac	ctgtggcatg	tttgtctaca	gtaagtgaat	attatgggca	2040
gtgggtgata	gagtggtggg	tttgggtgtg	taattttttt	tttaattttt	acagttttgt	2100
ggttttaaaga	attttgtatt	gtgatttttt	taaaagggtc	tgtgtctgaa	cctgagctgt	2160
agcccgagcc	agaaccggag	cctgcaagac	ctaccgcggc	tcttaaaatg	gcgcctgcta	2220
tcctgagacg	cccgacatca	cctgtgtcta	gagaatgcaa	tagtagtacg	gatagctgtg	2280
actccgggtc	ttctaacaca	cctcctgaga	tacaccgggt	ggtcccgtcg	tgccccatta	2340
aaccagttgc	cgtgagagtt	ggtgggctgc	gccaggctgt	ggaatgtatc	gaggacttgc	2400
ttacagagcc	tgggcaacct	ttggacttga	gctgtaaacg	cccaggcca	taagggtgta	2460
acctgtgatt	gcgtgtgtgg	ttacgcctt	tgtttgcgtg	atgagttgat	gtaagtttaa	2520
taaggggtga	gataatgttt	aacttgcatg	cgtgtttaa	tggggcgggg	cttaaagggt	2580
atataatgct	ccgtgggcta	atcttgggtt	catctgacct	catggaggct	tgggagtggt	2640
tgggaagattt	ttctgctgtg	cgtaacctgc	tggaacagag	ctctaacagt	acctcttggg	2700
tttggagggtt	tctgtggggc	tcaccccagg	caaagttagt	ctgcagaatt	aaggaggatt	2760
acaagtggga	attgaagag	cttttgaat	cctgtgtgta	gctgtttgat	tcctttgaatc	2820
tgggtcacca	ggcgcttttc	caagagaagg	tcacaaagac	tttggatttt	tcacacccgg	2880
ggcgcgctgc	ggctgtgtgt	gcttttttga	gttttataaa	ggataaatgg	agcgaagaaa	2940
cccatctgag	cgggggggtac	ctgctggatt	ttctggccat	gcactctgtg	agagcggttg	3000
tgagacacaa	gaatcgctcg	ctactgttgt	cttccgtccg	ccggcgata	ataccgacgg	3060
aggagcagca	gcagcagcag	gaggaagcca	ggcgcgcgcg	gcaggagcag	agcccatgga	3120
accgagagc	cggcctggac	cctcgggaat	gaatgttgta	caggtggctg	aactgtatcc	3180
agaactgaga	cgcattttga	caattacaga	ggatgggcag	gggctaaagg	gggtaaagg	3240
ggagcggggg	gcttgtgagg	ctacagagga	ggctaggaat	ctagctttta	gcttaatgac	3300
cagacaccgt	cctgagtgtg	ttacttttca	acagatcaag	gataattgct	ctaattgagct	3360
tgatctgctg	gcgcagaagt	attccataga	gcagctgacc	acttactggc	tgccagccagg	3420
ggatgatttt	gaggaggcta	ttagggtata	tgcaaagggt	gcacttaggc	cagattgcaa	3480
gtacaagatc	agcaaaactt	taaatatcag	gaattgttgc	tacatttctg	ggaacggggc	3540
cgagggtggg	atagatacgg	aggatagggt	ggcctttaga	tgtagcatga	taaatatgtg	3600
gcccgggggtg	cttggcatgg	acgggggtgt	tattatgaat	gtaagggtta	ctggcccaaa	3660
ttttagcggg	acggtttttc	tggccaatac	caaccttatc	ctacacgggt	taagcttcta	3720
tgggtttaac	aatacctgtg	tgggaagcct	gaccgatgta	aggggttcggg	gctgtgcctt	3780
ttactgtctg	tgggaagggg	tggtgtgtcg	ccccaaaagc	agggcttcaa	ttaagaaatg	3840
cctctttgaa	aggtgtacct	tgggtatcct	gtctgagggt	aactccaggg	tgccgccaaa	3900
tgtggcctcc	gactgtgggt	gcttcatgct	agtgaagagc	gtggctgtga	ttaagcataa	3960
catgggtatg	ggcaactgcg	aggacagggc	ctctcagatg	ctgacctgct	cggacggcaa	4020
ctgtcacctg	ctgaagacca	ttcacgtagc	cagccactct	cgcaaggcct	ggccagtggt	4080
tgagcataac	atactgacct	gctgttcctt	aacaggaggg	aacaggaggg	gggtgttctc	4140
accttaccaa	tgcaatttga	gtcacactaa	gatattgctt	gagcccgaga	gcattgtcca	4200
gggtgaacctg	aacgggggtg	ttgacatgac	catgaagatc	tgggaaggtg	tgaggtagca	4260
tgagaccggc	accaggtgca	gaccctgcga	gtgtggcggt	aaacatatta	ggaaccagcc	4320
tgtgatgctg	gatgtgaccg	aggagctgag	gcccgatcac	ttgggtgctg	cctgcacccg	4380
cgctgagttt	ggctctagcg	atgaagatac	agattgaggt	actgaaatgt	gtggcgctgg	4440

cttaaggggtg	ggaaagaata	tataaggtgg	gggtcttatg	tagttttgta	tctgttttgc	4500
agcagccgcc	gcccgcctga	gcaccaactc	gtttgatgga	agcattgtga	gctcatattt	4560
gacaacgcgc	atgcccccat	gggcccgggt	gcgtcagaat	gtgatgggct	ccagcattga	4620
tggtcgcccc	gtcctgcccc	caaaactctac	taccttgacc	tacgagaccg	tgtctggaac	4680
gccgtttggag	actgcagcct	ccgcgcgcgc	ttcagccgct	gcagccaccg	cccgcgggat	4740
tgtgactgac	tttgctttcc	tgagcccgcct	tgcaagcagt	gcagcttccc	gttcctccgc	4800
ccgcgatgac	aagttgacgg	ctcttttggc	acaattggat	tctttgaccc	gggaacttaa	4860
tgctgcttct	cagcagctgt	tggatctgcg	ccagcagggt	tctgccctga	aggcttcctc	4920
ccctcccaat	gcggttttaa	acataaataa	aaaaaccagc	tctgtttgga	tttgatcaa	4980
gcaagtgtct	tgctgtctct	cgagggatct	ttgtgaagga	accttacttc	tggtggtgta	5040
cataattgga	caaaactacct	acagagattt	aaagctctaa	ggtaaatata	aaatttttaa	5100
gtgtataatg	tggttaaacta	ctgattctaa	ttgtttgtgt	attttagatt	ccaacctatg	5160
gaactgatga	atgggagcag	tggtggaatg	cctttaatga	ggaaaacctg	ttttgctcag	5220
aagaaatgcc	atctagtgat	gatgaggcta	ctgctgactc	tcaacattct	actcctccaa	5280
aaaagaagag	aaaggtagaa	gaccccaagg	actttccttc	agaattgcta	agttttttga	5340
gtcatgctgt	gtttagtaat	agaactcttg	cttgctttgc	tatttacacc	acaaaggaaa	5400
aagctgcact	gctatacaag	aaaattatgg	cttcttactcc	tgtaaccttt	ataagtaggc	5460
ataacagtta	taatcataac	atactgtttt	ccttttagctt	acacaggcat	agagtgtctg	5520
ctattaataa	ctatgctcaa	aaattgtgta	cttagagatca	tttaatttgt	aaaggggtta	5580
ataaggaata	tttgatgtat	agtgccttga	ccacacctcc	taatcagcca	taccacattt	5640
gtagaggttt	tacttgcttt	aaaaaacctc	cctgaacctc	gaaacataaa	gaaacataaa	5700
atgaatgcaa	ttgttggtgt	taacttggtt	attgcagctt	ataatgggta	caaataaagc	5760
aatagcatca	caaatttcac	aaataaagca	tttttttcac	tgcatctctag	ttgtggtttg	5820
tccaaactca	tcaatgtatc	ttatcatgtc	tgatccggc	tgtggaatgt	gtgtcagtta	5880
gggtgtggaa	agtccccagg	ctccccagca	ggcagaagta	tgcaaagcat	gcatctcaat	5940
tagtcagcaa	ccaggtgtgg	aaagtccccca	ggctccccag	caggcagaag	tatgcaaagc	6000
atgcatctca	attagtcagc	aaccatagtc	ccgcccctaa	ctccgcccct	cccgcctcta	6060
actccgcccc	gttccgcccc	ttctccgccc	caggtgtgac	taattttttt	tatttttgca	6120
gaggccgagg	ccgcctcggc	ctctgagcta	ttccagaagt	agtgaggagg	cttttttgga	6180
ggcctaggct	tttgcaaaaa	gcttggacac	aagacaggct	tgcgagatat	gtttgagaat	6240
accactttat	cccgctcag	ggagaggcag	tgcgtaaaaa	gacgaggact	catgtgaaat	6300
actgggtttt	agtgcgcccag	atctctataa	tctcgcgcaa	cctattttcc	cctcgaaacac	6360
tttttaagcc	gtagataaac	aggctgggac	acttcacatg	agcgaaaaat	acatcgtcac	6420
ctgggacatg	ttgcagatcc	atgcacgtaa	actcgcaagc	cgactgatgc	cttctgaaca	6480
atggaaaggg	attattgccc	taagccgtgg	cgggtctgga	ccgggtgcgt	tactggcgcg	6540
tgaactgggt	attcgtcatg	tcgataccgt	ttgtatttcc	agctacgatc	acgacaacca	6600
gcgcgagctt	aaagtgtgta	aacgcgcaga	aggcgatggc	gaaggcttca	tcgttattga	6660
tgacctgggt	gataccgggt	gtactgcggt	tgcgattcgt	gaaatgtatc	caaaagcgca	6720
ctttgtcacc	atcttcgcaa	aaccggctgg	tcgtccgctg	gttgatgact	atgttgttga	6780
tatcccgcaa	gatacctgga	ttgaacagcc	gtgggatatg	ggcgctcgat	tcgtcccgcc	6840
aatctccggg	cgctaattctt	ttcaacgcct	ggcactgccc	ggcggtgttc	tttttaactt	6900
caggcggggt	acaatagttt	ccagtaagta	ttctggagge	tgcatccatg	acacaggcaa	6960
acctgagcga	aacctgttcc	aaaccccgct	ttaaaccatcc	tgaaacctcg	acgctagtcc	7020
gccgctttaa	tcacggcgca	caaccgcctg	tgcatcgcc	ccttgatggg	aaaaccatcc	7080
ctcactggta	tcgcatgatt	aaccgtctga	tggtgattctg	gcgcggcatt	gacccacgcg	7140
aaatcctcga	cgtccaggca	cgtattgtga	tgagcgatgc	cgaacgtacc	gacgatgatt	7200
tatacgatac	gggtgattggc	taccgtggcg	gcaactggat	ttatgagtgg	gccccggatc	7260
tttgtgaagg	aaccttactt	ctgtggtgtg	acataattgg	acaaactacc	tacagagatt	7320
taaagctcta	aggttaaata	aaaattttta	agtgtataat	gtgttaaact	actgattcta	7380
attgtttgtg	tatttttagat	tccaacctat	ggaactgatg	aatgggagca	gtggtggaat	7440
gcctttaatg	aggaaaacct	gttttgctca	gaagaaatgc	catctagtga	tgatgaggct	7500
actgctgact	ctcaacattc	tactcctcca	aaaaagaaga	gaaaggtaga	agaccccaag	7560
gactttcctt	cagaattgct	aagttttttg	agtcatgctg	tgtttagtaa	tagaactctt	7620
gcttgctttg	ctattttacac	cacaaaggaa	aaagctgcac	tgctatacaa	gaaaattatg	7680
gaaaaatatt	ctgtaacctt	tataagtagg	cataacagtt	ataatcataa	catactgttt	7740
tttcttactc	cacacaggca	tagagtgtct	gctattaata	actatgctca	aaaattgtgt	7800
acctttagct	ttttaatttg	taaaggggtt	aataaggaat	atttgatgta	tagtgccttg	7860
actagagatc	ataatcagcc	ataccacatt	tgtagagggt	ttacttgctt	taaaaacct	7920
cccacacctc	cccctgaacc	tgaaacataa	aatgaatgca	attgttgttg	ttacttgtt	7980
tattgcagct	tataatgggt	acaaataaag	caatagcatc	acaaatttca	caaataaagc	8040
atthttttca	ctgcattcta	gttggtggtt	gtccaaactc	atcaatgtat	cttatcatg	8100
ctggatcccc	aggaagctcc	tctgtgtcct	cataaaccct	aacctcctct	acttgaggag	8160

```

acattccaat cataggtgc ccatccaccc tctgtgtcct cctgttaatt aggtcactta 8220
acaaaaagga aattgggtag ggggtttttca cagaccgctt tctaagggtta attttaaaat 8280
atctgggaag tcccttccac tgctgtgttc cagaagtgtt ggtaaacagc ccacaaatgt 8340
caacagcaga aacatacaag ctgtcagctt tgcacaaggg cccaacaccc tgctcatcaa 8400
gaagcactgt ggttgctgtg ttagtaatgt gcaaaacagg aggcacattt tcccacctg 8460
tgtaggttcc aaaatatcta gtgttttcat ttttacttgg atcaggaacc cagcactcca 8520
ctggataagc attatcctta tccaaaacag ccttgtggtc agtgttcac tctgtactgt 8580
caactgtagc attttttggg gttacagttt gagcaggata tttggtcctg tagtttgcta 8640
acacaccctg cagctccaaa gggtcccccac caacagcaaa aaaatgaaaa tttgacctt 8700
gaatgggttt tccagcacca ttttcatgag tttttgtgt ccctgaatgc aagttaaaca 8760
tagcagttac cccaataacc tcagttttta cagtaacagc tccccacatc aaaatatttc 8820
cacaggttaa gtctcattt aaattaggca aaggaattct tgaagacgaa agggcctcgt 8880
gatacgcccta tttttatagg ttaatgtcat gataataatg gtttcttaga cgtcaggtgg 8940
cacttttcgg ggaaatgtgc gcggaacccc tatttgttta ttttctaaa tacattcaa 9000
tatgtatccg ctcatgagac aataaccctg ataatgtct caataatatt gaaaaaggaa 9060
gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cattttgcct 9120
tctgttttt gctcaccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg 9180
tgacagagtg ggttacatcg aactggatct caacagcggt caacagcggt agatccttg 9240
ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt 9300
atcccggtgt gacgcgggc aagagcaact cggtcgccgc atacactatt ctcagaatga 9360
cttggttgag tactcaccag tcacagaaaa gcactcttac gatggcatga cagtaagaga 9420
attatgcagt gctgccataa ccatgagtga taactctgcg gccaaactac tctgacaac 9480
gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg 9540
ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac 9600
gatgcctgca gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactc 9660
agcttcccgg caacaattaa tagactggat ggaggcggat aaagtgtcag gaccttct 9720
gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg 9780
gtctcggggt atcattgcag cactggggcc agatggtaag ccctcccgtg tcgtagttag 9840
ctacacgacg gggagtccag caactatgga tgaacgaaat agacagatcg ctgagatagg 9900
tgctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat 9960
tgatttaaaa cttcattttt aattttaaag gatctaggtg aagatccttt ttgataatct 10020
catgacaaaa atcccttaac gtgagttttt gttccactga gcgtcagacc ccgtagaaaa 10080
gatcaaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaacaaa 10140
aaaaccacgg taccagcgg tggtttgttt gccggatcaa gagctaccaa ctctttttcc 10200
gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta 10260
gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgtc tgctaattcc 10320
gttaccagtg gctgctgcca gtggcgataa gtctgtctt accgggttgg actcaagacg 10380
atagttaccg gataaggcgc agcggctcggg ctgaacgggg cgtgagctat cacagcccag 10440
cttgagcga acgacctaca aggcggacag gtatccggtg agcggcaggg gagaaagcgc 10500
cacgctccc gaaggagaa aggggagctt cagggggaaa cgtgagctat tcggaacagg 10560
agagcgcacg agggagctt ctcgattttt gtgatgctcg tcaggggggc ggagcctatg 10620
tcgccacctc tgacttgagc gctgattttt gttcctggcc ttttgctggc cttttgctca 10740
gaaaaacgcc agcaacgcgg cctttttacg cccctgatcc tgtggataac cgtattaccg 10800
catgttcttt cctgcgttat gccgaacgac cgagcgcagc gactcagtg gctgagtg 10860
agctgatacc gctcgcgca attttctcct tacgcactgt tgcggtattt cacaccgcat 10920
ggaagagcgc ctgagtcggt tctgctctga tgccgcatag ttaagccagt atacactcg 10980
atggtgcact ctcatgacaa catggctgcg ccccgacacc cgccaacacc cgctgacgcg 11040
ctatcgctac gtgactgggt cccggcatcc gcttacagac aagctgtgac cgtctccggg 11100
ccctgacggg cttgtctgct ttcaccgtca tcaccgaaac gcgcgaggca gc 11152
agctgcatgt gtcagaggtt

```

<210> 13

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 13

gacggatcgg gagatctcc

<210> 14
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 14
ccgcctcaga agccatagag cc

22

<210> 15
<211> 14455
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: plasmid

<400> 15
aagcttgggc agaaatgggtt gaactcccga gagtgtccta cacctagggg agaagcagcc 60
aaggggttgt tccccaccaa ggacgaccgc tctgcgcaca aacggatgag cccatcagac 120
aaagacatat tcattctctg ctgcaaaactt ggcatagctc tgctttgcct ggggctattg 180
ggggaagttg cggttcgtgc tcgcagggtc ctccaccctg actcttttaa tagctcttct 240
gtgcaagatt acaatctaaa caattcggag aactcgacct tcctcctgag gcaaggacca 300
cagccaactt cctcttataa gccgcacatga ttttgcctct cagaaataga aataagaatg 360
cttgctaaaa attatatattt taccaataag accaatccaa taggtagatt attagttact 420
atgttaagaa atgaatcatt atcttttagt actattttta ctcaaattca gaagttagaa 480
atgggaatag aaaatagaaa gagacgctca acctcaattg aagaacaggt gcaaggacta 540
ttgaccacag gcctagaagt aaaaaaggga aaaaagagtg tttttgtcaa aataggagac 600
aggtgggtggc aaccaggagc ttatagggga ccttacatct acagaccaac agatgcccc 660
ttaccatata caggaagata tgacttaaat tgggataggt ggggtacagt caatggctat 720
aaagtgttat atagatccct cccttttcgt gaaagactcg ccagagctag acctccttgg 780
tgtatgttgt ctcaagaaga aaaagacgac atgaaacaac aggtacatga ttatatattt 840
ctaggaacag gaatgcactt ttggggaaag attttccata ccaaggaggg gacagtggct 900
ggactaatag aacattattc tgcaaaaact catggcatga gttattatga atagccttta 960
ttggcccaac cttgcgggtc ccagggttta agtaagtttt tgggttcaaaa ctgttcttaa 1020
aacgaggatg tgagacaagt ggtttcctga cttgggttgg tatcaaaggt tctgatctga 1080
gctctgagtg ttctattttc ctatgttctt ttggaattta tccaaatctt atgtaaatgc 1140
ttatgtaaac caagatataa aagagtgtcg attttttgag taaacttgca acagtccctaa 1200
cattcacctc ttgtgtgttt gtgtctgttc gccatccctg ctccgctcgt cacttatcct 1260
tcactttcca gaggtcccc cctcggataa gtgacccttg tctctatttc tactatttgg 1320
gctggcgccc gaacagggac cctcggataa gtgacccttg tctctatttc tactatttgg 1380
tgtttgtctt gtattgtctc tttcttctct ggctatcatc acaagagcgg aacggactca 1440
ccatagggac caagctagcg actgaaaatg agacatatta tctgccacgg aggtgttatt 1500
accgaagaaa tggccgccag tcttttggac cagctgatcg aagaggtact ggctgataat 1560
cttccacctc cttagcattt tgaaccacct acccttcacg aactgtatga ttttagacgtg 1620
acggcccccg aagatcccaa cgaggaggcg gtttcgcaga tttttccga ctctgtaatg 1680
ttggcggtgc aggaagggat tgacttactc acttttccgc cggcgcccgg ttctcggag 1740
ccgcctcacc tttcccgga gcccgagcag cggagcaga gagccttggg tccggtttct 1800
atgccaaacc ttgtaccgga ggtgatcgat cttacctgcc acgaggctgg ctttccaccc 1860
agtacgacg aggatgaaga gggtagggag tttgtgttag attatgtgga gcaccccg 1920
cacgggttgc ggtcttgtca ttatcaccgg aggaatacgg gggaccaga tattatgtgt 1980
tcgctttgtc atataggac ctgtggcatg tttgtctaca gtaagtgaat attatgggca 2040
gtgggtgata gagtgggtggg tttgtgtgtg taattttttt ttttaattttt acagttttgg 2100
ggtttaaaaga attttgtatt gtgatttttt taaaagggtc tgtgtctgaa cctgagcctg 2160
agccccagcc agaaccggag cctgcaagac ctaccgcggc tcctaaaatg gcgcctgcta 2220
tccctgagc cccgacatca cctgtgtcta gagaatgcaa tagtagtacg gatagctgtg 2280
actccggtcc ttctaacaca cctcctgaga tacaccgggt ggtcccgctg tgccccatta 2340
aaccagttgc cgtgagagtt ggtggcgctc gccaggctgt ggaatgtatc gaggacttgc 2400
ttaacgagcc tgggcaacct ttggacttga gctgtaaacg ccccaggcca taaggtgtaa 2460

acctgtgatt gcgtgtgtgg ttaacgcctt tgtttgctga atgagttgat gtaagtttaa 2520
taaaggggtga gataatgttt aacttgcctg gcgtgttaaa tggggcgggg cttaaagggg 2580
atataatgcg ccgtgggcta atcttgggtta catctgacct catggaggct tgggagtggt 2640
tggaagatttt ttctgctgtg cgtaacttgc tgaacacagag ctctaacagt acctcttggt 2700
tttggagggtt tctgtggggc tcatcccagg caaagttagt ctgcagaatt aaggaggatt 2760
acaagtggga atttgaagag cttttgaaat cctgtgggtga gctgtttgat tctttgaatc 2820
tgggtcacca ggcgcttttc caagagaagg tcatcaagac tttggatttt tccacaccgg 2880
ggcgcgctgc ggctgctgtt gcttttttga gttttataaa ggataaatgg agcgaagaaa 2940
cccctctgag cgggggggtac ctgctggatt ttctggccat gcatctgtgg agagcgggtg 3000
tgagacacaa gaatcgctg ctactgttgt cttccgtccg cccggcgata ataccgacgg 3060
aggagcagca gcagcagcag gaggaagcca ggcggcggcg caggagcag aactgtatcc 3180
acccgagagc cggcctggac cctcgggaat gaattgttga cagggtggctg gggtaaaagag 3240
agaactgaga cgcatttttga caattacaga ggatgggcag gggctaaaagg gcttaatgac 3300
ggagcggggg gcttgtgagg ctacagagga ggctaggaaat ctactgttga gataattgag 3360
cagacaccgt cctgagtgta ttacttttca acagatcaag acttactggc tgcagccagg 3420
tgatctgctg gcgcagaagt attccataga gcagctgacc gcacttaggc cagattgcaa 3480
ggatgatttt gagggagcta ttaggtata tgcaaagggtg gacatttctg ggaacggggc 3540
gtacaagatc agcaaacttg taaatatcag gaattgttgc tacatttctg taaatatgtg 3600
cgagggtggag atagatacgg aggatagggt ggcctttaga tgtagcatga ctggcccaa 3660
gcccgggggtg cttggcatgg acgggggtgt tattatgaat caaccttatc ctacacgggtg 3720
ttttagcggg acgggttttcc tggccaatac caaccttatc aggggttcggg gctgtgcctt 3780
tgggtttaac aatacctgtg tgggaagcctg gaccgatgta aggggttcggg ttaagaaatg 3840
tactgctgc tggaaggggg tgggtgtctg ccccaaaaagc aggggttcgaa tgcgcccaa 3900
cctctttgaa aggtgtacct tgggtatcet gtctgagggt aactccaggg tgccgcacaa 3960
tgtggcctcc gactgtggtt gcttcatgct agtgaaaagc gtggctgtga cggacggcaa 4020
catggtatgt ggcaactgcg aggacagggc ctctcagatg cgcaaggcct ggccagtggt 4080
ctgtcacctg ctgaagacca ttcacgtagc gctgttccct aacaggaggg ggggtgttcc 4140
tgagcataac atactgacct gctgttccct gatattgctt gagcccgaga gcatgtccaa 4200
accttaccac tgcaatttga gtcacactaa catgaagatc tggaaaggtg tgaggtacga 4260
ggtgaacctg aacgggggtg ttgacatgac gtgtggcggg aaacatatta ggaaccagcc 4320
tgagaccgcg accaggtgca gacctgcga gctgtggcgg gcccgatcac ttgggtgctg cctgcaccg 4380
tgtgatgctg gatgtgaccg aggagctgag gcccgatcac actgaaatgt gggggcgtgg 4440
cgctgagttt ggctctagcg atgaagatac agattgaggt tagttttgta tctgttttgc 4500
cttaaggggt ggaaagaata tataaggtgg gggctctatg gtttgatgga agcattgtga gctcatattt 4560
agcagccgcc gccgccatga gggccggggg gcgtcagaat gtgatgggct ccagattga 4620
gacaacgcgc atgcccccat ggcctgcccg caaactctac taccttgacc tgcagaccg 4680
tggctgcccc gtccctgccc ccccgccgct ttcagccgct tgcaagcagt gcagcttccc 4740
gccgttggag actgcagcct ttgtctttcc tgagcccgtc acaattggat tctttgacct 4800
tgtgactgac aagttagcgg ctcttttggc tggatctgag cccagcgggt tctgtttgga 4860
ccctcccaat gcggtttaa acataaataa aaaccagac acccttactt 4920
gcaagtgtct tgctgtctct cagagatctt ttgtgaagga accttactt 5040
cataattgga caaactacct acagagattt aaagctctaa ggtaaatata aaatttttaa 5100
gtgtataatg tgtaaaacta ctgattctaa ttgtttgtgt atttttagatt ccaacctatg 5160
gaactgatga atgggagcag tgggtggaatg cctttaatga ggaaaacctg tttgtctcag 5220
aagaaatgcc atctagtgt gatgaggcta ctgctgactc tcaacattct actcctccaa 5280
aaaagaagag aaaggtagaa gaccccaagg actttccttc ttttacacc acaaaggaaa 5340
gtcatgctgt gtttagtaat agaactcttg cttgctttgc tatttacacc ataagtaggc 5400
aagctgcact gctatacaag aaaattatgg aaaaatattc tgtaaccttt ataagtgctg 5460
ataacagtta taatcataac atactgtttt ttcttactcc ttttaattgt aaaggggtta 5520
ctattaataa ctatgctcaa aaattgtgta ctagagatca taatcagcca taccacattt 5580
ataaggaata ttgatgtat agtgccctga ccacacctcc cctgaacct ataagtaggc 5640
gtagagggtt tacttgcttt aaaaaacctc attgcagctt ataaggtta gaaacataaa 5700
atgaatgcaa ttgttgttgt taacttgttt tttttttcac tgcattctag ttgtgggttg 5760
aatagcatca caaatttcac aaataaagca tttttttcac tgtggaatgt gtgtcagtta 5820
tccaaactca tcaatgtatc ttatcatgtc tggatccggc tggcagaagta caggcagaag 5880
gggtgtggaa agtccccagg ctccccagca ggcctcccag ctccgcccat cccgccctta 5940
tagtcagcaa ccaggtgtgg aaagtcccca cccgccctaa ctccgcccat cccgccctta 6000
atgcatctca attagtcagc aaccatagtc catggctgac taattttttt tatttatgca 6060
actccgccca gttccgccca ttctccgcc ctctgagcta ttccagaagt agtgaggagg ctttttttga 6120
gaggccgagg ccgcctcggc ctctgagcta ttccagaagt agtgaggagg ctttttttga 6180

ggcctaggct tttgcaaaaa gcttggacac aagacaggct tgcgagatat gtttgagaat 6240
accactttat cccgcgtcag ggagaggcag tgcgtaaaaa gacgcggact catgtgaaat 6300
actgggtttt agtgcgccag atctctataa tctcgcgcaa cctattttcc cctcgaacac 6360
tttttaagcc gtagataaac aggtctgggac acttcacatg agcgaaaaat acatcgtcac 6420
ctgggacatg ttgcagatcc atgcacgtaa actcgcaagc cgactgatgc cttctgaaca 6480
atggaaagggc attattgccg taagccgtgg cggctctggta cccgggtgcgt tactggcgcg 6540
tgaactgggt attcgtcatg tcgataccgt ttgtatttcc agctacgatc acgacaacca 6600
gcgcgagcgt aaagtgcgtg aacgcgcaga aggcgatggc gaaggcttca tcgttattga 6660
tgacctgggt gataccgggt gtactgcggt tgcgattcgt gttgatgact atgttggtga 6720
ctttgtcacc atcttcgcaa aaccggctgg tcgtccgctg gtcgggatag ggcgtcgtat 6780
tatcccgcaa gatacctgga ttgaacagcc gtgggatag ggcgtcgtat ttttaccctt 6840
aatctccggt cgctaattct ttcaacgcct ggcactgccg ggcgttggtc ttttaactt 6900
caggcgggtt acaatagttt ccagtaagta ttctggaggc tgcattccatg acacaggcaa 6960
acctgagcga aacctgttc aaaccccgct ttaaacatcc tgaaacctcg acgctagtcc 7020
gccgctttaa tcacggcgca caaccgcctg tgcagtcggc ccttgatggg gacccacgcg 7140
ctcactggta tcgcatgatt aaccgtctga tbtggatctg gcgcggcatt gaccgatgatt 7200
aaatccctga cgtccaggca cgtattgtga tgcagatgc cgaacgtacc gacgatgatt 7260
tatacgatac ggtgattggc tactgtggcg gcaactggat ttatgagtgg gcccggatg 7320
tttgtgaagg aacctactt ctgtgggtgtg acataattgg agtgtataat gtgttaaact actgattcta 7380
taaagctcta aggtaaatat aaaattttta ggaactgatg aatgggagca gtggtggaat 7440
attgtttggg tatttttagat tccaacctat gtagatctg gaagaaatgc catctagtga tgatgaggct 7500
gcctttaatg aggaaaacct gttttgctca gaagaaatgc catctagtga tgatgaggct 7560
actgctgact ctcaacattc tactctcca aaaaagaaga gaaaggtaga agacccaag 7620
gactttcctt cagaattgct cacaaggaa aaagctgcac tgctatacaa gaaaattatg 7680
gcttgccttg ctattttacac tataagtagg cataacagtt ataatacaa catactgttt 7740
gaaaaatatt ctgtaacctt tagagtgtct gctattaata actatgctca aaaattgtgt 7800
tttcttactc cacacaggca taaaggggtt aataaggaat atttgatgta tagtgccctg 7860
acctttagct ttttaatttg taaaggggtt tgtagagggt ttacttgctt taaaaaacct 7920
actagagatc ataatacagc ataccacatt tgtagaggtt attgttggtt taaacttgtt 7980
cccacacctc cccctgaacc tgaaacataa aatgaatgca acaaatttca caaataaagc 8040
tattgcagct tataatgggt acaataaag gtccaaactc atcaatgtat cttatcatgt 8100
atttttttca ctgcattcta gttgtgggtt cataaaccct aacctcctct acttgagagg 8160
ctggatcccc aggaagctcc tctgtgtcct cctgtgtcct tcttaagggt tcttaagggt 8220
acattccaat cataggctgc ccatccaccc cagacggctt ggttaaacagc ccacaaatgt 8340
acaaaaagga aattgggtag ggggttttca cagaagtgtt ggttaaacagc ccacaaatgt 8400
atctgggaag tcccttccac ctgtcagcgt ttagtaatgt ggttaaacagc ccacaaatgt 8460
caacagcaga aacatacaag gtgttttcat ttttacttgg atcaggaacc cagcactcca 8520
gaagcactgt ggttgctgtg ttttacttgg ccttgggtgc gagcaggata tttgggtcctg tagtttgcta 8580
tgtaggttcc aaaatatcta tccaaacag gttacagttt caacagcaaa aaaatgaaaa tttgaccctt 8640
ctggataagc attatcctta gtttccctg ggttccccac ttttcatgag ccttgaatgc aagtttaaca 8700
caactgtagc attttttggg gtttccctg ttttcatgag ccttgaatgc aagtttaaca 8760
acacaccctg cagctccaaa ggttccccac ttttcatgag ccttgaatgc aagtttaaca 8820
gaatgggttt tccagcacca tccagcacca ttttcatgag ccttgaatgc aagtttaaca 8880
tagcagttac cccaataacc tcaatagga aaggaattct tgaagacgaa agggcctcgt 8940
cacagggtta gtccctcatt ttaatgtcat gataataatg gtttcttaga cgtcaggtgg 9000
gatacgctta tttttatagg gcggaacccc tatttgttta tttttctaaa tacattcaaa 9060
cacttttccg ggaatgtgc aataaccctg ataaatgctt caataatatt ttttttgcgt 9120
tatgtatccg ctcatgagac tccgtgtcgc ccttattccc gatgctgaag atcagttggg 9180
gagtatgagt attcaacatt aaacgctggg gaaagtaaaa aagatccttg agagttttcg 9240
tctgtttttt gctcaccag aactggatct caacagcggt tttttaaagt ctgctatgtg gcgcgggtatt 9300
tgcacgagtg ggttacatcg tcatgagcac tttttaaagt atacactatt ctgagatga 9360
ccccgaagaa cgttttccaa aagagcaact cggctgcggc gatggcatga cagtaagaga 9420
atcccggtgt gacgcggggc tcacagaaaa gcatcttacg gatggcatga cagtaagaga 9480
cttggttgag tactcaccag ccactgagtg taacactgcg gccaacttac atgttaactc 9540
attatgcagt gctgccataa taaccgcttt ttgtcacac aacgacgagc gtgacaccac 9600
gatcggagga ccgaaggagc agctgaatga agcatacca actggcgaac tacttactct 9660
ccttgatcgt tgggaaccgg caacgttgcg caaactatta aaagtgcag gaccacttct 9720
gatgcctgca gcaatggcaa tagactggat ggagggcgat tctggagccg gtgagcgtgg 9780
agcttcccg cttccggctg gctgggttat agatggtaag cctcccgta tcgtagttat 9840
gcgctcggcc atcattgcag cactggggcc caactatgga tgaacgaat agacagatcg ctgagatagg 9900
ctacacgacg gggagtcagg caactatgga tgaacgaat agacagatcg ctgagatagg 9900

tgccctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tacttttagat 9960
tgatttataaaa cttcattttt aatttataaag gatctaggtg aagatcccttt ttgataatct 10020
catgaccaaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 10080
gatcaaaagga tcttcttgag atccctttttt tctgcgcgta atctgctgct tgcaaacaaa 10140
aaaaccaccg ctaccagcgg tgggtttgtt gccggatcaa gagctaccaa ctctttttcc 10200
gaaggttaact ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta 10260
gttagggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaattcct 10320
gttaccagtgt gctgctgcca gtggcgataa gtcgtgtott accgggttgg actcaagacg 10380
atagttaccg gataaggcgc agcggtcggg ctgaacgggg gggtcgtgca cacagcccag 10440
cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc 10500
cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg ctgtatagtc 10560
agagcgcacg agggagcttc cagggggaaa cgcctgggtat ctttatagtc ctgtcgggtt 10620
tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagccctatg 10680
gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc cttttgtctg 10740
catgttcttt cctgcgttat cccctgatcc tgtggataac cgtattaccg cctttgagtg 10800
agctgatacc gctcgcgcga gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc 10860
ggaagagcgc ctgatgcggg attttctcct tacgcatctg tgccggtattt cacaccgcat 10920
accgcctcag aagccataga gccaccgcga tccccagcat gcctgctatt gtcttcccaa 10980
tcctccccct tgctgtcctg cccaccacca cccccagaa cctactcaga 11040
caatgctagt caatttctc attttattag gaaaggacag tgggagtgcc accttccagg 11100
gtcaaggaag gcacggggga ggggcaaaaca acagatggct ggcaactaga aggcacagtc 11160
gaggctgatc agcagagctct agcatttagg tgacactata gaatagggccc ctctagatgc 11220
atgctcgagc ggccgcttct ttattcttgg gcaatgtatg gcaatgtatg agaggatgtg 11280
gcaaatattt cattaatgta gttgtggcca gaccagtccc atgaaaatga catagagtat 11340
gcacttgagc ttgtgtctcc tgtttcctgt gtaccgttta gtgtaattgg tagtggttaca 11400
ggtttagttt tgtctccggt taagtaaaact tgactgacaa tgttactttt ggcagtttta 11460
ccgtgagatt ttggataagc tgataggtta ggcataaaatc caacagcgtt tgtataggct 11520
gtgccttcag taagatctcc atttctaaag ttccaatatt ctgggtccag gaaggaattg 11580
tttagtagca ctccattttc gtcaaatctt ataataagat gagcactttg aactgttcca 11640
gatattggag ccaaactgcc tttaacagcc gcatctttct ctgtagcaag tctactgta 11700
ccacattttt ttaagaccaa agtgagttta agcttatcat tatttttgtt tctactgta 11760
ggagatggag ctgggtgtgg ccacaaagtt agcttctagt taggaaccat agccttggtt 11820
atggcacctg tgctgtcaaa actaaggcca gttcctagtt gatttggtgt ttggtgcatta 11880
gaatcaaatt ctaggccatg gccaattttt ttatgggcta tggctgtagc tggtgcatta 11940
gggtgaaccaa attcaagccc atctcctgca ttaatgggta tggctgtagc gtcaaacatc 12000
aacccttggc cagtgccttag gttaacctca agctttttgg aattgtttga agctgtaaac 12060
aagtaaaggc ctttgttgta gttaatatcc aagttgtggg ctgagtttat aaaaagaggg 12120
ccctgtccta gtcttagatt tagttggttt tgagcatcaa acggataact aacatcaagt 12180
ataagcgctc tgttttgaga atcaatcctt agtcctcctg ctacattaag ttgcatattg 12240
ccttgtgaat caaaacccaa ggctccagta actttagttt gcaaggaagt attattaata 12300
gtcacacctg gaccagttgc tacgggtcaaa gtgttttagt cgtctgttac atgcaaagga 12360
gccccgtact ttagtcctag ttttccattt tgtgtataaa tgggctcttt caagtcaatg 12420
cccaagctac cagtggcagt agtttagagg ggtgaggcag tgatagtaag ggtactgcta 12480
tcggtgggtg tgagggggcc tgatgtttgc agggctagct tctcttctga cactgtgagg 12540
ggctccttgg tggtcaatgct aagtttggag tctgtcacgg ttagcggggc ctgtgattgc 12600
atggtgagtg tgttgccgc gaccattaga ggtgcccggc cagccacagt tagggcttct 12660
gaggttaact tgaggggtgc agatatttcc aggtttatgt ttgacttggg ttttttgaga 12720
ggtgggctca cagtggttac attttgggag gtaagggttc cggcctcgtc cagagagagg 12780
ccgttgccca ttttgagcgc aagcatgcc aagtggttaa cttagaggtc ggataggcgc 12840
aaagagagta cccaggggg actctcttga aacccattgg gggatacaaa gggaggagta 12900
agaaaaggca cagttggagg accgggttcc gtgtcatatg gatacacggg gttgaaggt 12960
tcttcagacg gtcttgccgc cttcatcttg gatctcaagc ctgccacacc tcacctcgac 13020
catccgcccgt ctcaagaccg cctactttaa ttacatcatc agcagacct ccccgagaaa 13080
caaccccagc cgccaccgc tgccgcccgc caggtgtgct agcctacctt gcgactgtga 13140
ctggtttagc gcctttctcg agagggtttc cgatccggtc gatcgaggc tactggaaag 13200
cctcggtggc ggagtaccgt tcggaggccg acgggtttcc gatccaagag gatccaagag 13260
accgcgaaga gtttgcctc aaccgcgagc ccaacagcga gctcgaattc agatccgagc 13320
tcggtagcaa gcttgggtct ccctatagtg agtcgtatta atttcgataa gccagtaagc 13380
agtgggttct ctagttagcc agagagctct gcttatatag acctcccacc gtacacgctc 13440
accgcccatt tgcgtcaatg ggccggagtt gttacgacat tttggaaagt cccgttgatt 13500
ttggtagcaa aacaaactcc cattgacgtc aatgggtgg agacttgga atccccgtga 13560
gtcaaacccg tatccacgcc cattgatgta ctgcaaaac cgcatacaca tggtaatagc 13620


```

gatgactaat acgtagatgt actgccaagt aggaaagtcc cataaggtca tgtactgggc 13680
ataatgccag gcggggccatt taccgtcatt gacgtcaata gggggcgtagc ttggcatatg 13740
atacacttga tgtactggcca agtgggcagt ttaccgtaaa tagtccaccc attgacgtca 13800
atggaaagtc cctattggcg ttactatggg aacatacgtc attattgacg tcaatgggcg 13860
ggggtcgttg ggcggtcagc caggcgggcc atttacgta agttatgtaa cgcggaactc 13920
catatatggg ctatgaacta atgaccccg tgcggaagc agcgcaaac gcctaaccct 14040
tcaatgtcaa cgcgatatatc tggcccgtag gctcggtcaa gccttgccct gttgtagctt 14100
aagcagattc ttcattgcaat tgcggtcaa gacagggagc agatactggc ttaactatgc 14160
cgcgactac tcagcgacct ccaacacaca agcagggagc tggggagatc tccgatccg 14220
ggcatcagag cagattgtac tgagagtcga ccatagggga tggggagatc agtataact 14280
tctatgggtg actctcagta caatctgctc tgatggcgca tagttaagcc acccgctgac 14340
ccgctatcgc tacgtgactg ggtcatggct gcgccccgac acccgccaac acccgctgac 14400
gcgcctgac gggcttgtct gctcccgcga tccgcttaca gacaagctgt gaccgtctcc 14455
gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag gcagc

```

<210> 16

<211> 10610

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid

<400> 16

```

gacggatcgg gagatccgag cggtacacag aattcaggag acacaactcc aagtgcatac 60
tctatgtcat tttcatggga ctggtctggc cacaactaca ttaatgaaat atttgccaca 120
tcctcttaca ctttttcata cattgcccac gaataaagaa tcggttgtgt tatgtttcaa 180
cgtgtttatt tttcaattgc agaaaatttc aagtcatttt tcattcagta gtatagcccc 240
accacacat agcttatata gatcacgta ccttaatcaa actcacagaa ccctagtatt 300
caacctgcca cctccctccc aacacacaga gtacacagtc ctttctcccc ggctggcctt 360
aaaaagcatc atatcatggg taacagacat attccttaggt gttatattcc acacggtttc 420
ctgtcgagcc aaacgctcat cagtcatatt aataaactcc ccgggcagct cacttaagtt 480
catctcgtcg tccagctgct gagccacagg ctgctgtcca acttgcggtt gcttaacggg 540
cgcggaagga gaagtcacg cctacatggg ggtagagtca taatcgtgca tcaggatagg 600
gcggtgggtg tcagcagcg cggaataaaa ctgctgccgc cgccgctccg tcctgcagga 660
atacaacatg gcagtggtct cctcagcgat gattcgacc gcccgagca taaggcgct 720
tgctctccgg gcacagcagc gcaccctgat ctacttaaa tcagcacagt aactgcagca 780
cagcaccaca atattgttca aaatcccaca gtgcaaggcg ctgtatccaa agctcatggc 840
ggggaccaca gaaccacagt ggccatcata ccacaagcgc aggtagatta agtggcgacc 900
cctcataaac acgctggaca taaacattac ctcttttggc atgttgtaat tcaccacctc 960
cgggtaccat ataaacctct gattaaacat ggcccatcc accaccatcc taaaccagct 1020
ggccaaaacc tgcccgcggg ctatacactg cagggaaacc ggactggaac aatgacagt 1080
gagagccag gactcgtaac catggatcat catgctcgtc atgatatcaa tgttggcaca 1140
acacagggac acgtgcatac acttcctcag gattacaagc tcctcccggg ttagaacctc 1200
atcccaggga acaaccatt cctgaatcag cgtaaatccc acactgcagg gaagacctcg 1260
cacgtaactc acgttgtgca ttgtcaaagt gttacattcg ggcagcagcg gatgatctc 1320
cagtatggta gcgcggggtt ctgtctcaaa aggaggtaga cgatccctac tgtacggagt 1380
gcgcggagac aaccgagatc gtgttggtcg tagtgtcatg ccaaattggaa cgccggagct 1440
agtcataatt cctgaagcaa aaccagggtg gggcgtaga aacagatctg cgtctccggg 1500
ctcgccgctt agatcgctct tctatgtaaa ctccttcagc cgccgctgcc ctgataacat 1620
ggcgccccct ggcttcgggt tctatgtaaa acaccagcc aacctacaca ttcgttctgc gactcacaca 1680
ccaccaccgc agaataagcc acaccagcc aacctacaca tttattccaa aagattatcc 1740
cgggaggagc ggggaagagt tattaagtga acgcgctccc ctccggtggc gtggtcaaac 1800
aaaacctcaa aatgaagatc aatggcattt gtaagatgtt gcacaatggc ttccaaaagg 1860
tctacagcca aagaacagat aatggcattt gtaagatgtt cttcaggggt aatctcctct 1920
caaacggccc tcagtcocaa gtggacgtaa aggctaaacc cttcaggggt aatctcctct 1980
ataaacattc cagcaccttc aacctgccc aaataattct cctctcaat 2040
atatctctaa gcaaatccg aatattaagt ccggccattg taaaaatctg ctccagagcg 2100
ccctccacct tcagcctcaa gcagcgaatc atgattgcaa aaattcaggt tcctcacaga 2160
cctgtataag attcaaaagc ggaacattaa caaaaatacc gcgatcccg aggtcccttc 2220
gcagggccag ctgaacataa tcgtgcaggt ctgcacggac cagcgcgggc acttccccc 2280

```

caggaacctt gacaaaagaa cccacactga ttatgacacg cataactcgga gctatgctaa 2280
ccagcgtagc cccgatgtaa gctttgttgc atggggcgcg atataaaatg caaggtgctg 2340
ctcaaaaaat caggcaaaagc ctcgcgcaaa aaagaaagca catcgtagtc atgctcatgc 2400
agataaaaggc aggtaaagctc cggaaaccacc acagaaaaag acaccatttt tctctcaaac 2460
atgtctgcgg gtttctgcat aaacacaaaa taaaataaca aaaaaacatt taaacattag 2520
aagcctgtct tacaacagga aaaacaaccc ttataagcat aagacggact acggccatgc 2580
cggcgtgacc gtaaaaaaac tggtcaccgt gattaaaaag caccaccgac agctcctcgc 2640
tcatgtccgg agtcataatg taagactcgg taaacacatc aggttgattc atcggtcagt 2700
gctaaaaagc gaccgaaata gcccggggga atacataccc gcaggcgtag agacaacatt 2760
acagccccc a taggaggtat aacaaaatta ataggagaga aaaacacata aacacctgaa 2820
aaaccctcct gcctaggcaa aatagcacc tcccgcctca gaacaacata cagcgcttca 2880
cagcggcagc ctaacagtc a gccttaccag taaaaaagaa aacctattaa aaaaacacca 2940
ctcgacacgg caccagctca atcagtcaca gtgtaaaaaa gggccaagtg cagagcgagt 3000
atataatgga ctaaaaaatg acgtaacggt taaagtccac aaaaaacacc cagaaaaaccg 3060
cacgcgaacc tacgcccaga aacgaaagcc aaaaaaccca caacttcctc tggctcagtc 3120
ttccgttttc ccacgttacg taacttcccg gatcctctcc cgatccctta tggctcagtc 3180
tcagtacaat ctgctctgat gccgcatagt taagccagta tctgtctcct gcttctgtgt 3240
tggaggctgc tgagttagtc gcgagcaaaa ttaagctac aacaaggcaa ggcttgaccg 3300
acaattgcat gaagaatctg cttaggggta ggcttttgc gctgcttcgc gatgtacggg 3360
ccagatatac gcgttgacat tgattattga ctagtattta atagtaatca attacgggg 3420
cattagtcca tagcccatat atggagttcc gcgttacata acttacggta aatggccgc 3480
ctggctgacc gcccaacgac cccgcgccat tgacgtcaat aatgacgtat gttcccatag 3540
taacgccaat agggactttc cattgacgtc aatgggtgga ccctattgac gtaaatgacg 3600
acttggcagt acatcaagtgc atgtacgccc caagtacgccc atgggacttt cctacttggc 3660
gtaaatggcc gcctggcat tatgcccagt acatgacctt atgggacttt cctacttggc 3720
agtaacatcta cgtattagtc atcgtatata ccatgggtgat gcgggttttg cagtacatca 3780
atgggcgtgg atagcgggtt gactcacggg gatttccaag tctccacccc attgacgtca 3840
atgggagttt gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactccg 3900
cccattgac gcaaatgggc ggtaggcggtg tacgggtggga ggtctatata agcagagctc 3960
tctggctaac tagagaaccc actgcttact ggcttatcga aattaatacg actcactata 4020
gggagaccca agcttggtac cgagctcgga gtactcttgg atcggaaccc cgtcgccctc 4080
ggtgaggaca aactcttcgc agggacctga gcgagtcggc atcgaccgga ctcggaaaacc 4140
tctcgagaaa ggcgtctaac cagtcacagt cggaaggtag gctgagcacc gctgagcacc 4200
gcagcgggtg gcggtcgggg ttgtttctgg cggaggtgct gctgatgatg taattaaagt 4260
aggcgggtct gagacggcg atgggtcgagg tgaggtgtgg caggcttgag atccaagatg 4320
aagcgcgcaa gaccgtctga agataccttc aaccccggtg atcccatatga cagcgaaacc 4380
ggtcctccaa ctgtgccttt tcttactcct cctttgtat ccccaatgg gtttcaagag 4440
agtccccctg gggtaacttc tttgcgccta tccgaacctc tagttacctc caatggcatg 4500
cttgcgctca aaatgggcaa cggcctctct ctggacgagg ccggcaacct tacctcccaa 4560
aatgtaacca ctgtgagccc acctctcaaa aaaaccaagt ccaacataaa cctggaaata 4620
tctgcacccc tcacagttac ctcagaagcc ctaactgtgg ctgcccgcgc acctctaatt 4680
gtcgcgggca acacactcac catgcaatca caggccccgc taaccgtgca cgactccaaa 4740
cttagcattg ccacccaagg accctcaca gtgtcagaag gaaagctagc cctgcaaaaca 4800
tcaggccccc tcaccaccac ccatagcagt acccttacta tcaactgcct cctgcaaaaca 4860
actactgcca ctggtagctt gggcattgac ttgaaaagag ccatttatac acaaaaatgga 4920
aaactaggac taaagtacgg ggctcctttg catgtaacag acgacctaaa cactttgacc 4980
gtagcaactg gtccaggtgt gactattaat aatacttctt tgcaaaactaa agttactgga 5040
gccttggggt ttgattcaca aggcaatatg caacttaatt tagcaggagg actaaggatt 5100
gattctcaaa acagacgcct tatacttgat gttagtatac cgtttgatgc tcaaaaacca 5160
ctaaatctaa gactaggaca gggccctctt tttataaact cagcccacaa cttggatatt 5220
aactacaaca aaggccttta cttgtttaca gtttcaaaac attocaaaaa gcttgagggt 5280
aacctaagca ctgccaaagg gttgatgttt gacgctacag ccatagccat taatgcagga 5340
gatgggcttg aatttgggtc acctaattgca ccaaacacaa atcccccaa aacaaaaatt 5400
ggccatggcc tagaatttga ttcaaacaaag gctatgggtc ctaaactagg aactggcctt 5460
agttttgaca gcacaggtgc cattacagta ggaacacaaa ataagtataa gctaactttg 5520
tggaaccacac cagctccatc toctaactgt agactaaatg cagagaaaga tgctaaactc 5580
actttgggtc taacaaaatg tggcagtc aaactttgct atacttgcta cagtttcagt tttggctgtt 5640
aaaggcagtt tggctccaat atctggaaca gttcaaagtg ctcactttat tataagattt 5700
gacgaaaatg gagtgctact aaacaattcc ttctggacc cagaatattg gaactttaga 5760
aatggagatc ttactgaagg cacagcctat acaaacgctg ttggatttat gcctaacct 5820
tcagcttatc caaaatctca cggtaaaact gccaaaagta acattgtcag tcaagtttac 5880

ttaaaccggag aaaaaactaa acctgtaaca ctaaccatta cactaaacgg tacacaggaa 6000
acaggagaca caactccaag tgcataactct atgtcatttt catgggactg gtctggccac 6060
aactacatta atgaaatatt tgccacatcc tcttacactt tttcatacat tgcccaagaa 6120
taaagaagcg gccgctcgag catgcatcta gagggcccta ttctatagtg tcacctaaat 6180
gctagagctc gctgatcagc ctgcactgtg ccttctagtt gccagccatc tgttgtttgc 6240
ccctcccccg tgcccttcctt gaccttgga ggtgccactc ccactgtcct ttctaataa 6300
aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg ggggtggggtg 6360
gggcaggaca gcaaggggga ggattggga gacaatagca ggcatgctgg ggatggcggtg 6420
ggctctatgg cttctgaggc ggaaagaacc agctggggct ctagggggta tccccacgg 6480
ccctgtagcg gcgcattaaag cgccggcggt gtgggtggta cgcgacgct gaccgctaca 6540
cttgccagcg ccctagcgcc cgctccttcc gcttcttcc cttccttct cgccacgttc 6600
gccggcttcc ccgctcaagc tctaaatcgg ggcatccct taggggttccg atttagtgc 6660
ttacggcacc tcgaccccaa aaaacttgat taggggtgat gttcacgtag tgggccatcg 6720
ccctgataga cgggttttctg ccctttgacg tggagtcca cgttctttaa tagtgactc 6780
ttgttccaaa ctggaacaac actcaacct atctcggtct attcttttga tttataaggg 6840
atlttgggga tttcggccta ttggttaaaa aatgagctga tttacaaaa atttaacggc 6900
aattaattct gtggaatgtg tgtcagttag ggtgtggaaa gtccccaggc tccccaggca 6960
ggcagaagta tgcaaaagcat gcactcaat tagtcagcaa ccaggtgtgg aaagtcccca 7020
ggctccccag caggcagaag tatgcaaaagc atgcatctca attagttagc aaccatagtc 7080
ccgccccctaa ctccgccccat cccgccccct actccgccc gttccgccc ttctccgcc 7140
catggctgac taattttttt tatttatgca gaggcggagg cgcctctgc ctctgagcta 7200
ttccagaagt agtgaggagg ctttttttga ggcctaggct tttgcaaaaa gctcccgga 7260
gcttgtatat gattgcacgc atctgatcaa gagacaggat gaggatcgtt tcgcatgatt 7320
gaacaagatg aacagacaat aggttctccg gccgcttggg tggagaggct attcggctat 7380
gactgggcac ccatcttccg cggctgctct gatgcccg cctgaatga actgcaggac 7440
gggcgccccg ttctttttgt gctggccacg acgggcgttc cttgcccagc tgtgctcgac 7500
gaggcagcgc ggctatcgtg gactggctg ctattgggag aagtgcggg gcaggatctc 7560
gttgtcactg aagcgggaag ggactggctg ctattgggag tggctgatgc aatgcggcg 7620
ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcg 7680
ctgcatacgc ttgatccggc tacctgccc atcgaccacc aagcgaaaca tcgcatcgag 7740
cgagcacgta ctcggtatgga agccggctct gtgatcagg atgatctgga cgaagagcat 7800
caggggctcg cgccagccga actgttccg aggtcaagg cgcgatgcc cgacggcgag 7860
gatctcgctg tgacccatgg cgtatgcctg ttgccgaata tcatgggtgga aaatggccgc 7920
tttctggat ttcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg 7980
ttgggtaccc gtgatattgc tgaagagctt ggcggcgaaat gggctgaccg cttcctcgtg 8040
ctttacggta tcgcccgtcc cgattcgag cgcacgcct agcagcgccc aacctgccat 8100
ttcttctgag cgggactctg gggttcgaaa tagccgacca agcagcgccc aacctgttcc 8160
cacgagattt cgattccacc gccgccttct atgaaagggt gggcttcgga atcgttttcc 8220
gggacgcgg ctggatgatc ctccagcgcg gggatctcat caatagcatc acaaatttca 8280
ccaacttgtt tattgcagct tataatgggt acaaataaag gtccaaactc atcaatgtat 8340
caataaagc atttttttca ctgcattcta gttgtgggt gtccaaactc atcaatgtat 8400
cttatcatgt ctgtataccg tcgacctcta gctagagctt ggcgtaatca tggctatagc 8460
tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca 8520
taaagtgtaa agcctggggt gcctaattgag tgagctaact cacattaatt gcggttgcgt 8580
cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcgccaac 8640
gcgcggggag aggcgggttg cgtattgggc gctcttccgc ttctcgtc actgactcgc 8700
tgccgctcgt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgg 8760
tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg 8820
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg 8880
agcatcaca aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 8940
accaggcggt tccccctgga agctccctcg tgccgtctcc tgttccgacc ctgccgtta 9000
cgggataact gtccgccttt ctcccttcgg gaagcgtggc gctttctcaa tgcacagct 9060
gtaggatatc cagttcggtg taggtcggtc gctccaagct ggtgtgtg gggctgtgtg 9120
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtc aacctggtaa 9180
gacacgact atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 9240
taggcgggtg tacagagttc ttgaagtggg ggctaacta cggctacact agaaggacag 9300
tatttgggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 9360
gatccggcaa acaaaccacc gctggtagcg gtgggttttt tgttgcaag cagcagatta 9420
cgcgcaaaaa aaaaggatct caagaagatc ctttgatctt ttctacggg tctgacgctc 9480
agtggaaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 9540
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 9600
cttgggtctga cagttacca tgcttaatca gtgaggcacc tatctcagcg atctgtctat 9660

ttcgttcac	catagttgcc	tgactccccg	tcgtgtagat	aactacgata	cgggagggct	9720
taccatctgg	ccccagtgc	gcaatgatac	cgcgagaccc	acgctcaccc	gctccagatt	9780
tatcagcaat	aaaccagcca	gccggaaggg	ccgagcgcag	aagtggctcc	gcaactttat	9840
ccgcctccat	ccagtctatt	aattggtgcc	gggaagctag	agtaagtagt	tcgccagtta	9900
atagtttgcg	caacgttggt	gccattgcta	cagggcatcgt	ggtgtcacgc	tcgtcgtttg	9960
gtatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	tcccccatgt	10020
tgtgcaaaaa	agcggttagc	tccttcgggtc	ctccgatcgt	tgtcagaagt	aagttggccg	10080
cagtgttatc	actcatgggt	atggcagcac	tgcataattc	tcttactgtc	atgccatccg	10140
taagatgctt	ttctgtgact	ggtgagtact	caaccaagtc	attctgagaa	tagtgtatgc	10200
ggcgaccgag	ttgctcttgc	ccggcgctcaa	tacgggataa	taccgcgcca	catagcagaa	10260
ctttaaaagt	gctcatcatt	ggaaaaacgtt	cttcggggcg	aaaactctca	aggatcttac	10320
cgctgttgag	atccagttcg	atgtaaccca	ctcgtgcacc	caactgatct	tcagcatctt	10380
ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaaatgcc	gcaaaaaagg	10440
gaataagggc	gacacggaaa	tggtgaatac	tcatactctt	cctttttcaa	tattattgaa	10500
gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	10560
aacaaatagg	ggttcgcgcg	acatttcccc	gaaaagtgcc	acctgacgtc		10610

<210> 17
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Primer

<400> 17
 tgtacaccgg atccggcgca cacc

24

<210> 18
 <211> 35
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Primer

<400> 18
 cacaacgagc tcaattaatt aattgccaca tcttc

35

<210> 19
 <211> 4
 <212> PRT
 <213> adenovirus

<400> 19
 Thr Leu Trp Thr
 1

<210> 20
 <211> 12
 <212> PRT
 <213> adenovirus

<400> 20
 Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly Ser
 1 5 10

<210> 22
 <211> 327
 <212> DNA

<213> adenovirus

<400> 22
agatctgaat tcgagctcgc tgttgggctc gcggttgagg acaaactctt cgcggtcttt 60
ccagtactct tggatcgga acccgctggc ctccgaacgg tactccgcca ccgagggacc 120
tgagcgagtc cgcacgcacc ggatcgga aaactctcgag aaaggcgtct aaccagtcac 180
agtcgcaagg taggctgagc accgtggcgg gcggcagcgg gtggcggtcg gggttgtttc 240
tggcgagggt gctgctgatg atgtaattaa agtaggcggg cttgagacgg cggatggtcg 300
aggtagggtg tggcaggctt gagatct 327

<210> 23

<211> 32480

<212> DNA

<213> adenovirus

<400> 23
catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt 60
ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtggt gcggaagtgt 120
gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180
gtgtgcgcgg gtgtacacag gaagtgaaca ttttcgcgcg gttttaggcg gatgtttag 240
taaatttggg cgtaaccgag taagatttgg ccattttcgc gggaaaactg aataagagga 300
agtgaatct gaataatttt gtgttactca tagcgcgtaa tctctagcat cgatgtcgac 360
aagcttgaat tcgattaatg tgagttagct cactcattag gcaccccagg ctttacactt 420
tatgcttccg gctcgtatgt tgtgtggaat tgtgagcggg taacaatttc acacaggaaa 480
cagctatgac catgattacg aattcggcgc agcaccatgg cctgaaataa cctctgaaag 540
aggaacttgg ttaggtacct tctgaggcgg aaagaaccag ctgtggaatg tgtgtcagtt 600
agggtgtgga aagtcgcccg gctccccagc aggcagaagt atgcaaagca tgcactctca 660
ttagtcagca accaggtgtg gaaagtcccc aggtccccca gcaggcagaa gtatgcaaag 720
catgcatctc aattagtcag caaccatagt cccgcccta actccgccca tcccgccctt 780
aactccgccc agtccgccc cctctgagct attccagaag tagtgaggag gcttttttgg 900
agaggccgag gccgcctcgg cctctgagct attccagaag tagtgaggag gcttttttgg 960
aggcctaggc ttttgcaaaa agcttgggat ctctataatc tcgcgcaacc tattttcccc 1020
tcgaacactt ttttaagccgt agataaacag gctgggacac ttcacatgag cgaaaaatac 1080
atcgtaacct gggacatggt gcagatccat gcacgtaaac tcgcaagccg actgatgcct 1140
tctgaacaat ggaaaggcat tattgccgta agccgtggcg gtctgggtacc ggtgggtgaa 1200
gaccagaaac agcacctcga actgagccgc gatattgccc agcgtttcaa cgcgctgtat 1260
ggcgagatcg atcccgtcg tttacaacgt cgtgactggg aaaacctgg cgttacccaa 1320
cttaategce ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggcccg 1380
accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgctt tgccctgggt 1440
ccggcaccag aagcggtgcc ggaaagctgg ctggagtgcg atcttcctga ggccgatac 1500
gtcgtcgtcc cctcaaacg gcagatgcac ggttacgatg cgcccatcta caccaacgta 1560
acctatccca ttacggtcaa tccgcgcttt gttcccacgg agaatccgac ggggtgttac 1620
tcgctcacat ttaatggtga tgaaagctgg ctacaggaag gccagacgag aattattttt 1680
gatggcggtta actcggcggt tcatctgtgg tgcaacgggc gctgggtcgg ttacggccag 1740
gacagtcggt tgccgtctga atttgacctg agcgcatttt tacgcgccgg agaaaaccgc 1800
ctcgcggtga tgggtgctgc ttggagtgc ggcagttatc tggaaatca ggatatgtgg 1860
cggtatgagc gcattttcgg tgacgtctcg ttgctgcata aaccgactac acaaatcagc 1920
gattttccatg ttgccactcg ctttaatgat gatttcagcc gcgctgtact ggaggctgaa 1980
gttcagatgt gcggcgagtt gcgtgactac ctacgggtaa cagtttcttt atggcagggt 2040
gaaacgcagg tcgccagcgg caccgcgcct ttccggcggt aaattatcga tgagcgtggt 2100
ggttatgccg atcgcgtcac actacgtctg aacgtcgaaa acccgaaact gtggagcgcc 2160
gaaatcccg atctctatcg tgcgggtggt gaactgcaca ccgccgacgg cacgctgatt 2220
gaagcagaag cctgcgatgt cggtttcgcg gaggtgcgga ttgaaaatgg tctgctgctg 2280
ctgaacggca agccgttgct gattcagagg gtttaaccgtc acgagcatca tctctgcat 2340
ggtcaggtea tggatgagca gacgatggtg caggatatcc tgctgatgaa gcagaacaac 2400
tttaacggcg tgcgtgttc gcattatccg aaccatccgc tgtggtacac gctgtgcgac 2460
cgctacggcc tgtatgtggt ggatgaagcc aatattgaaa cccacggcat ggtgccaatg 2520
aatcgtctga ccgatgatcc gcgtgggcta ccggcgatga gcgaacgcgt aacgcgaatg 2580
gtgcagcgcg atcgtaatca cccgagtgtg atcatctggt cgctggggaa tgaatcaggc 2640
cacggcgcta atcacgcgc gctgtatcgc tggatcaaat ctgtcgatcc tccccgccg 2700
gtgcagtatg aaggcgcgcg agccgacacc acggccaccg atattatttg cccgatgtac 2760
gcgcgcgtgg atgaagacca gcccttcccc gctgtgccga aatgggtccat caaaaaatgg

ctttcgctac ctggagagac ggcggcgctg atccttttgcg aatacgccca cgcgatgggt 2820
aacagtcttg ggcgttttcg taaatactgg caggcgtttc gtcagtatcc ccgtttacag 2880
ggcgacttcg tctgggactg ggtggatcag tcgctgatta aatatgatga aaacggcaac 2940
ccgtggctcg cttacggcgg tgattttggc gatagccga acgatcgcca gttctgtatg 3000
aacggtcttg tcttttccga ccgcacgcgg catccagcgc tgacggaagc aaaacaccag 3060
cagcagtttt tccagttccg tttatccggg caaaccatcg aagtgaccag cgaatacctg 3120
ttccgtcata gcgataacga gctcctgcac tggatgggtg cgctggatgg taagccgctg 3180
gcaagcggtg aagtgcctct ggatgtcgt ccacaaggta aacagttgat tgaactgcct 3240
gaactaccgc agccggagag cgcggggcaa ctctggctca cagtacgct agtgcaaccg 3300
aacgcgaccg catggtcaga agccgggcac atcagcgctt ggcagcagtg gcgtctggcg 3360
gaaaacctca gtgtgacgt ccccgccggg tcccacgcca tcccgcatct gaccaccagc 3420
gaaatggatt tttgcatcga gctgggtaat aagcgttggc aatttaaccg ccagtacggc 3480
tttctttcac agatgtggat tggcgataaa aaacaactgc tgacgcgct gcgcgatcag 3540
ttcaccgctg caccgctgga taacgacatt ggcgtaagtg aagcgaccgg cattgaccct 3600
aacgcctggg tcgaacgctg gaaggcgggc ggccattacc aggcogaagc agcgttgggtg 3660
cagtgcacgg cagatacact tgctgatgct gtgctgatta cgaccgctca cgcgtggcag 3720
catcagggga aaaccttatt tatcagccgg ggtgctgatta ggattgatgg tagtggtcaa 3780
atggcgatta ccgttgatgt tgaagtggcg agcgatacac cgcacccggc gcggattggc 3840
ctgaaactgcc agctggcgca ggtagcagag cgggtaaaact ggctcggatt agggccgcaa 3900
gaaaactatc ccgaccgctt tactgcccgc tgttttgacc gctgggatct gccattgtca 3960
gacatgtata cccgtacgt cttcccgagc gaataacggc tgcgtgcgg gagcgcgcaa 4020
ttgaattatg gccacacca gtggcgcggc gacttccagt tcaacatcag ccgctacagt 4080
caacagcaac tgatggaaac cagccatcgc catctgctgc acgcggaaga aggcacatgg 4140
ctgaatatcg acggtttcca tatggggatt ggtggcgacg actcctggag cccgtcagta 4200
tcggcggaat tccagctgag cgcgggtcgc taccattacc agttggtctg gtgtcaaaaa 4260
taataataac cggcgaggcc atgtctgccc gtatttcgcg taaggaaatc cattatgtac 4320
tatttaaaaa acacaaactt ttggatgttc ggtttattct tttctttta cttttttatc 4380
atgggagcct acttcccgtt tttcccgtat tctctgttct tggtacatg acatcaacca tatcagcaa 4440
agtgtacgg gtattatatt tgcgctatt tctctgttct cgctattatt ccaaccgctg 4500
tttggcttgc tttctgacaa actcggaaact tctctgttct agcttataat ggttacaaat 4560
aaagcaatag catcacaat ttcacaaata aagcattttt ttcactgcat ttagttgtg 4620
gtttgtccaa actcatcaat gtatcttacc atgtctggat ccagatctgg gcgtggctta 4680
agggtgggaa agaataata aggtgggggt cttatgtagt tttgtatctg ttttgagca 4740
gccgcgccc ccatgagcac caactcgttt gatggaagca ttgtgagctc atatttgaca 4800
acgcgcgatg ccccatgggc cgggggtgct cagaatgtga tgggctccag cattgatgg 4860
cgccccgtcc tgccccgaaa ctctactacc ttgacctacg agaccgtgtc tggaacgcgc 4920
ttggagactg cagcctccgc cgcgcttca gccgtgcag ccaccgccc cgggatttg 4980
actgactttg ctttctgag cccgcttgca agcagtgag cttcccgttc atccgcccgc 5040
gatgacaagt tgacggctct tttggcaca ttggattctt tgaccggga acttaatgtc 5100
gtttctcagc agctgttggg tctgcgccag caggtttctg ccctgaaggc ttcctcccct 5160
cccaatgcgg tttaaaacat aaataaaaaa ccagactctg tttggatttg gatcaagcaa 5220
gtgtcttgct gtctttatct aggggttttg cgcgcgcggg agggccggga ccagcggtct 5280
cggtcgttga gggctcgtg ttttttttcc agaacgtggg aaagggtgact ctggatgttc 5340
agatacatgg gcataagccc gtctctgggg tggaggtagc accactgcag agcttcatgc 5400
tgccgggttg tgtgttagat gatccagtc tagcaggagc gctgggctg gtgcctaaaa 5460
atgtctttca gtacgaagc gattgccagg ggcaggccct tgggtgaagt gtttacaag 5520
cggttaagct gggatgggtg catacgtggg gatattagat gcactctgga ctgtattttt 5580
aggttggcta tgttcccagc catatccctc cggggattca tgttgtgcag aaccaccagc 5640
acagtgtatc cgggtgactt gggaaatttg tcatgtagct tagaaggaaa tgcgtggaag 5700
aacttgga ga cgccttgtg acctccaaga ttttccatgc attcgtccat aatgatggca 5760
atgggcccac gggcgcggc atagggcatt tttacaaagc gatcactaac gtcatagttg 5820
tgttccagga tgagatcgtc atagggcatt tttacaaagc gcgggcgag ggtgccagac 5880
tgccgtataa tgggtccatc cggcccgagg gcgtagttac cctcacagat ttgcatttcc 5940
cacgctttga gttcagatgg ggggatcatg tctacctgcg gggcgatgaa gaaaacgggt 6000
tccggggtag gggagatcag ctgggaagaa agcagggttc tgagcagctg cgacttaccg 6060
cagccgggtg gcccgtaaat cacacctatt accgggtgca actggtagt aagagagctg 6120
cagctgccgt catccctgag cagggggggc acttctgtaa gcatgtccct gactcgcatg 6180
tttccctga ccaaatccgc tttgagaccg tccgcccag gcgatagcag ttcttgcaag 6240
gaagcaaaagt ttttcaaccg tttgagaccg gtcacctgct ctacggcatc tgcgtggaag 6300
ccaagcagtt ccaggcggtc tttcgtgta cggcagtagt cgtgctcgt 6360
atatctctc gtttcgcggg ttggggcggc cctcgtcagc gtagtctggg 6420
ccagacgggc cagggtcatg tctttccag ggcgcagggt 6480

tcacgggtgaa ggggtgcgct ccgggctgcg cgctggccag ggtgcgcttg aggctgggtcc 6540
tgctgggtgct gaagcgctgc cgggtcttcgc cctgcgcgtc ggccaggtag catttgacca 6600
tggtgtcata gtccagcccc tccgcggcgt ggccttggc gcgcagcttg cccttgagg 6660
aggcgccgca cgaggggcag tgcagacttt tgagggcgta gagcttgggc gcgagaaata 6720
ccgattccgg ggagtaggca tccgcgccgc agggccccga gacgggtctcg cattccacga 6780
gccagggtgag ctctggccgt tcgggggtcaa aaaccagggt tccccatgc tttttgatgc 6840
gtttcttacc tctgggtttcc atgagccggt cctcgagcgt ggtgacgaaa aggctgtccg 6900
tgtccccgta tacagacttg agaggcctgt cctcgagcgg tgttccgcgg tcctcctcgt 6960
atagaaactc ggaccactct gagacaaaagg ctgcgcgtcca ggccagcacg aaggaggcta 7020
agtgggaggg gtagcgggtcg ttgtccacta ggggggtccac tgcgtccagg gtgtgaagac 7080
acatgtcgcc ctcttcggca tcaagggaagg tgattggttt gtaggtgtag gccacgtgac 7140
cgggtgttcc tgaagggggg ctataaaaagg ggggtggggg gcgttcgtcc tcaactctctt 7200
ccgcatcgct gtctgcgagg gccagctggt ggggtgagta ctccctctga aaagcgggca 7260
tgacttctgc gctaagattg tcagtttcca aaaacgagga ggatttgata ttcacctggc 7320
ccgcggtgat gcctttgagg gtggcgcgat ccatctggtc agaaaagaca atctttttgt 7380
tgtcaagctt ggtggcaaac gaccctgtaga gggcgttggg cagcaacttg gcgatggagc 7440
gcagggtttg gtttttgtcg cgatcggcgc gctccttggc cgcatgttt agctgcacgt 7500
attcgcgccg aacgcaccgc cattcgggaa agacgggtgg gctcgtcgcg ggcaccagg 7560
gcacgcgcca acccggttg tgcagggtga caaggctaac gctgggtggc acctctccg 7620
gtaggcgctc gttggtccag cagaggcggc cgcccttgcg cgagcagaat ggcggtaggg 7680
ggctctagctg cgtctcgtcc ggggggtctg cgtccacggt aaagacccc ggacgacggc 7740
gcgcgtcgaa gtagtctatc ttgcatcctt gcaagctag cgctgctgc catgcgcggg 7800
cggcaagcgc cgcctcgat ggggtgagtg ggggaccca tggcatgggg tgggtgagcg 7860
cggaggcgta catgccgcaa atgtcgtaaa cgtagagggg ctctctgagt attccaagat 7920
atgtagggtg gcatcttcca ccgcgtagtc tggcgcgcac gtaatcgat agttcgtgcg 7980
agggagcgag gaggtcggga ccgagggtgc tggcgcgag ctgctctgct cggaagacta 8040
tctgcttgaa gatggcatgt gagttggatg atatggttgg acgctggag acgttgaagc 8100
tggcgctgtg gagacctacc gcgtcacgca cgaaggaggc gtaggagtcg cgcagcttgt 8160
tgaccagctc ggcgggtgacc tgcacgtcta acagctcgcg gttcagggtt tccttgatga 8220
tgtcatactt atcctgtccc ttttttttcc cgtcggcctc cgaacggtaa aactcttcgc 8280
ggtctttcca gtactcttg atcgaaacc agcatccctt ggtgtccctg gagcctagca 8340
tgtagaactg gttgacggcc tggtaggcgc gagggtgagg tagcgcaaaa agcgctatg 8400
cctgcgcggc cttccggagc gaggtgtggg tggcgcaaaa ggtgtccctg accatgactt 8460
tgaggtaactg gtattgaaag tcagtgctgt cgcacccgcc ctgctcccag agcaaaaagt 8520
ccgtgcgctt tttggaacgc ggatttggca gggcgaaagg gacatcgtt aagagtatct 8580
ttcccgcgcg aggcataaag ttgctgtgta tccgggaagg gttgatgtt tggccacaa 8640
tgtaattac ctgggcggcg agcacgatc cgtcaaagcc gttgatgtt tggccacaa 8700
tgtaagttc caagaagcgc gggatgccct tgatggaagg caatttttta agttcctcgt 8760
agggtgagctc ttcaggggag ctgagcccg gctctgaaag gggccagctc gcaagatgag 8820
ggttggaagc gacgaatgag ctccacagg caggggccat tagcatttgc aggtgggtcgc 8880
gaaaggtcct aaactggcga cctatggcca ttttttctg ggtgatgcag tagaaggtaa 8940
gggggtcttg ttccagcgg tcccatccaa ggttcgcggc taggtctcgc gcggcagta 9000
ctagaggctc atctccgccc aacttcatga ccagcatgaa gggcacgagc tcggtgcgag 9060
aggcccccct ccaagtatag gtctctacat cgtaggtgac aaagagacgc tgggtattga 9120
gatgcgagcc gatcgggaag aactggatct cccgccacca attggaggag tggctattga 9180
tgtggtgaaa gtagaagtc ctgacgagg gacgaactc gtgctggctt ttgtaaaaac 9240
gtgcgcagta ctggcagcgg tgcacgggct gttacatcct tggcggggtt ggctggtggt 9300
cgcgacaaag gaagcagagt ggggaatttga gcccctcgcc tggcggggtt agggggagtt 9360
cttctacttc ggctgcttgt ccttgaccgt ctggctgctc gagggggagtt acgggtggatc 9420
ggaccaccac gccgcgcgag cccaaagtcc agatgtccgc gcgcgcggt cgagcctga 9480
tgacaacatc gcgcagatgg gagctgtcca tggcttgagg ctcccgcggc gtcaggtcag 9540
gcgggagctc ctgcaggttt acctcgcata gacgggtcag ggcgcgggct agatccaggt 9600
gatacctaatt ttccaggggc tgggttggtgg cggcgctcgat ggccttgaag aggcgcgac 9660
cccgcggcgc gactacggta ccgcgcggcg ggcgggtggc cgcggggggtg tccttggatg 9720
atgcatctaa aagcggtgac gcgggcgagc ccccgagggt aggggggggt cggacccgc 9780
cgggagaggg ggcaggggca cgtcggcgcc gcgcgcgggc agggagctgg gctgcgcgcg 9840
taggttgctg gcgaacgcga cgacgcggcg gttgatctcc tgaatctggc gcctctgcgt 9900
gaagacgacg gggccgggtg gcttgagcct gaaagagagt tgcacagaat caatttcggt 9960
gtcgttgacg gcggcctggc gcaaaatctc ctgacgtct cctgagttgt cttgatagtc 10020
gatctcgccc atgaactgct cgatctcttc ctcttgaga tctccgcgtc cggtcgcgtc 10080
cacgggtggc gcgaggtcgt tggaaatgag ggccatgagc tgcgagaagg cgttgaggcc 10140
tccctcgttc cagacgcggc tgtagaccac gccccttcg gcatcgcggg cgcgcatgac 10200

cacctgcegcg agattgagct ccacgtgccg ggcgaagacg gcgtagtttc gcaggcgctg 10260
aaagaggtag ttgagggttg tggcgggttg ttctgccacg aagaagtaca taaccacgcg 10320
tcgcaacgtg gattcggtga tatcccccac ggccctcaagg cgctccatgg cctcgtagaa 10380
gtccacggcg aagttgaaaa actgggagtt ggcgcgccgac acggttaact cctcctccag 10440
aagacggatg agctcgccga cagtgtccgc cacctcgccg tcaaaggcta cagggggcctc 10500
ttcttcttct tcaatctcct ctccataag ggccctccct tcttcttctt ctggcggcgg 10560
tgggggaggg gggacacggc ggcgacgacg ggcacccggg agggcggtcga caaagcgctc 10620
gatcatctcc ccgcggcgac ggcgcatggt ctccggtgacg gcgcggccgt tctcgccggg 10680
gcgcagttgg aagacccgc ccgtcatgtc ccggttatgg gttggcgggg ggctgccatg 10740
cggcagggat acggcgctaa ccatgcatct caacaattgt tgtgtaggta ctccgcggcc 10800
gagggacctg agcgagtccg catcgaccgg atcggaatac ctctcgagaa aggcgtctaa 10860
ccagtcacag tcgcaaggta ggctgagcac cgtggcgggc ggcagcgggc ggcggtcggg 10920
gttgtttctg gcggaggtgc tgctgatgat gtaattaaag taggcggtct tgagacggcg 10980
gatggtcgac agaagcacca tgtccttggg tccggcctgc tagtagtctt gcggtcgccg 11040
catgccccag gcttcgtttt gacatcgccg caggctcttg tagtagtctt gcatgagcct 11100
ttctaccggc acttcttctt ctcttctctc ttgtcctgca tctcttgcat ctatcgctgc 11160
ggcggcggcg gagtttggcc gtaggtggcg cctcttctct cccatgcgtg tgaccccgaa 11220
gccccctcat ggctgaagca gggctaggtc ggcgacaacg cgctcggcta atatggcctg 11280
ctgcacctgc gtgagggtag actggaagtc atccatgtcc acaaagcggg ggtatgcgcc 11340
cgtgttgatg gtgtaagtgc agttggccat aacggaccag ttaacggctc ggtgaccccg 11400
ctgcgagagc tccggtgtacc tgagacgcga gtaagccctc gagtcaaata cgtagtcgtt 11460
gcaagtcgcg accaggtact ggtatccac caaaaagtgc ggcggcggct ggcggtagag 11520
gggccagcgt aggggtggccg gggctccggg ggcgagatct tccaacataa ggcgatgata 11580
tccgtagatg tacctggaca tccaggtgat gccggcggcg gtggtggagg cgcgcggaaa 11640
gtcgcggacg aggttccaga tgttgccgag cggcaaaaag tgctccatgg tcgggacgtc 11700
ctggccggtc aggcgcgcgc aatcgttgac gctctagacc gtgcaaaaag agagcctgta 11760
agcgggcact cttccgtggg ctggtggata aatcgcaag ggtatcatgg cggacgaccg 11820
gggttccgag cccgtatccg gccgtccgcc gtgatccatg cggttaccgc ccgcgtgtcg 11880
aaccacgggtg tgcgacgtca gacaacgggg tgtgtctct tttggcttcc ttccaggcgc 11940
ggcggtcgtc gcgctagctt ttttggccac tggccgcgcg cagcgtaacg ggttaggctg 12000
gaaagcgaag gcattaagtg gctcgctccc tgtagccgga gggttatttt ccaaggggtg 12060
agtccgggga cccccggttc gagtctcgga ccggccggac tgcggcgaaac ggggggttgc 12120
ctccccgtca tcgaagacc cgtttgcaaa tctctccgga aacaggggac agcccccttt 12180
ttgcttttcc cagatgcac cgggtgctgc gcagatgcgc cccctctctc agcagcggca 12240
agagcaagag cagcggcaga catgcagggc accctccctc cctcctaccg cgtcaggagg 12300
ggcgacatcc gcggttgacg cggcagcaga ttggtgattac gaaccccgcc ggcgcgggc 12360
ccggcactac ctggacttgg aggagggcga gggcctggcg cggctaggag cgccctctcc 12420
tgagcggtag ccaagggtgc agctgaagcg tgatacgcgt gaggcgtacg tgcccgggca 12480
gaacctgttt cgcgaccgcg agggagagga gcccgaggag atgcgggatc gaaagtcca 12540
cgcagggcgc gagctgcggc atggcctgaa tcgcgagcgg ttgctgcgcg agggaggactt 12600
tgagcccgac gcgcgaaccg ggatttagtc cgcgcgcgca cagctggcgg ccgcccacct 12660
ggtaaccgca tacgagcaga cggtaacca ggagattaac tttcaaaaaa gctttaacaa 12720
ccacgtgcgt acgcttgtgg cgcgcgagga ggtggctata ggactgatgc atctgtggga 12780
ctttgtaagc cgcctggagc aaaacccaaa tagcaagccg ctcatggcgc agctgttctt 12840
tatagtgcag cacagcaggg acaacgaggc attcagggat gcgctgctaa acatagtaga 12900
gcccaggggc cgtgggtgc tcgatttgat aaacatctg cagagcatag tgggtcgagg 12960
gcgcagcttg agcctggctg acaaggtggc cgccatcaac tattccatgc ttagcctggg 13020
caagttttac gcccgcaaga tataccatac cccttacgtt cccatagaca aggaggtaaa 13080
gatcgagggg ttctacatgc gcatggcgct gaaggtgctt accttgagcg acgacctggg 13140
cgtttatcgc aacgagcgca tccacaaggc cgtgagcgtg agccggcggc gcgagctcag 13200
cgaccgcgag ctgatgcaca gacctgaaag ggcctggct tgacctgcgc tgggccccaa 13260
agaggccgag tcctactttg acgcgggcgc ggcggtggca cccgcgcgcg ggcacgggca 13320
cctggaggga gctggggccg gacctgggt ggagtagcag ccagaggacg ggcgagtag 13380
cggcggcggt gaggaatat acgaggacga tgagtaacgg acccgggcgt gcgagtacta 13440
agcggtagtg agccgtccgg ccttaactcc acggacgact ggcgcaggt gggggcggcg 13500
ctgcagagcc tgactgcgcg caatcctgac gcgttccggc agcagccgca ggccaaccgg 13560
atcatgtcgc ttctggaagc ggtggtccc ggcgcgcgca accccacgca ggcggcggct 13620
ctctccgcaa ttctggaagc ggtggtccc ggcgcgcgca accccacgca ggcggcggct 13680
ctggcgatcg taaacgcgct ggcggaatac agggccatcc ggcccgacga ggccggcctg 13740
gtctacgacg cgctgcttca gcgcgtggct cgttacaaca ggcgcaacgt gcagaccaac 13800
ctggaccggc tgggtggggg tgtgcgcgag gccgtggcgc agcgtgagcg cgcgcagcag 13860
cagggaacc tgggctccat ggttgacta aacgccttc tgagtacaca gcccggcaac 13920

gtgccgcggtgacagggagga ctacaccaac tttgtgagcg cactgcggtt aatggtgact 13980
gagacaccgc aaagtgaggt gtaccagtct gggccagact attttttcca gaccagtaga 14040
caaggcctgc agaccgtaaa cctgagccag gctttcaaaa acttgccagg gctgtggggg 14100
gtgcccgtgc ccacaggcga cgcgcgcacc gtgtctagct tgctgacgcc caactcgcgc 14160
ctgtttgtgc tgctaatagc gcccttcacg gacagtggca gcgtgtcccg ggacacatac 14220
ctaggtcact tgctgacact gtaccgcgag gccataggtc aggcgcattg ggacgagcat 14280
actttccagg agattacaag tgtcagccgc cgcgtggggc aggaggacac gggcagcctg 14340
gaggcaaccc taaactacct gtgaccaac cggcggcaga agatccctc gttgcacagt 14400
ttaaaccagc aggaggagcg ctttttgccg tacgtgcagc agagcgtgag ccttaacctg 14460
atgcccgcagc gggtaacgcc cagcgtggcg ctggacatga cgcgcgcgaa catggaaccg 14520
ggcatgtatg cctcaaaccg gccgtttatc aaccgcctaa tggactactt gcatcgcgcg 14580
gcccgcgtga cccccgagta tttaccaaat gccatcttga acccgactg gctaccgccc 14640
cctggtttct acaccggggg attcagagtg cccgagggta acgatggatt cctctgggac 14700
gacatagacg acagcgtggt gcgaaaggaa ccgcagacc tgctagagtt gcaacagcgc 14760
gagcagcgag aggcggcgct cggccccgcg gtcagatgct agtagcccat ttccaagctt gatagggtct 14880
ctaggcgctg ctcgcaccac ccgcccgcgc ctgctggcg aggaggagta cctaaacaac 14940
cttaccagca agccgcagcg cgaaaaaac ctgctccgg catttccaa caacgggata 15000
tcgctgctgc agccgcagcg gaggtagatg aagacgtacg cgcaggagca cagggacgtg 15060
gagagcctag tggacaagat ccgtcgtcaa aggcacgacc gtccagcggg tctggtgtgg 15120
ccaggcccgcc gcccgcaccac cgtcgtcaa ggcacgacc tgggaggag tggcaaccgc 15180
gaggacgatg actcggcaga cgacagcagc gtcctggatt tgggaggag catgatgcaa 15240
tttgccgacc ttcgccccag gctggggaga atgttttaa aaaaaaaaag cctcttagta 15300
aataaaaaaac tcaccaaggc catggcaccg agcgtttggt tctttgtatt ccccttagta 15360
tgccgcccgc ggcgatgtat gaggaaggtc ctcctccctc ctacgagagt gtggtgagcg 15420
cgccgcagct ggcggcgcg cctacgtcg cctacgggg ggagaaacag catccgttac tggcagcctg 15480
tgccctccgc gacacccacc cgtgtgtacc tgggtggacaa caagtcaacg gatgtggcat 15540
cacccttatt ccagaacgac acacagacca tctatcttga tcaatcttga aacaatgact 15600
ccctgaacta ggaggcaagc acacagacca tgccaaatgt gaacgagttc cactggggcg 15660
acagcccggg ggcagcagc caccatcctg atggtttacca atggtttacca 15720
gcgacctgaa aacctcctg atggtgtcgc gcttgccctac taaggacaat cagggtggagc 15780
ataagtttaa ggcgcgggtg ttcacgctgc cctagcctac ctactccgag accatgacca 15840
tgaaatacga gtgggtggag ttacgctgc cagaggggcaa agtgggcaga cagaacgggg 15900
tagaccttat gaacaacgcg atcgtggagc actacttgaa cttcagactg ggggttgacc 15960
ttctggaaag cgacatcggt gtaagtttg cctggggtat atacaaacga agccttccat ccagacatca 16020
ccgtcactgg aggatgcccgg gtggacttca cccacagccg cctgagcaac ttgttgggca 16080
ttttgtgccc gcaacccttc caggagggtc ttaggatcac ctacgatgat ctggagggtg 16140
tccgcaagcg cgcactgttg gatgtggacg cctaccaggc gagcttgaaa gatgacaccg 16200
gtaacattcc aacaggcgcg ggggtggcgca agcagcagtg cagcggcgcg gaagagaact 16260
ccaacgcggc agccgcggca atgcagccg tggaggacat ggccgaagca ggcggccgaag 16320
gcgacacctt tgccacacgg gctgaggaga agcgcgctga ggcgaagca gtgatcaaac 16380
ctgcccgcgc cgctgcgcaa cccgaggtcg agaagcctca gaagaaaccg gtgatcaaac 16440
ccctgacaga ggacagcaag aaacgcagtt acaacctaat aagcaatgac agcaccttca 16500
cccagtagac cagctgggtac cttgcataca actacggcga ccctcagacc ggaatccgct 16560
catggaccct gctttgcaat cctgacgtaa cctgcggctc ggagcaggtc tactgggtcgt 16620
tgccagacat gatgcaagac cccgtgacct tccgctccac ggcgcagatc agcaactttc 16680
cgggtgggtgg gcgcgagctg ttgcccgctg actccaagag cttctacaac gaccaggccg 16740
tctactccca actcatccgc cagtttacct ctctgaccca cgtgttcaat cgctttcccg 16800
agaaccagat tttggcgcg cccgagccc ccaccatcac caccgtcagt gaaaacggtc 16860
ctgctctcac agatcacggg acgctaccgc tgcgcaacag catcgaggga gtccagcgag 16920
tgaccattac tgacgcccaga cgcgcacct tttgagcaag gatgtttggc ggggccaaga 16980
tctcgcccg cgtcctatcg agccgcactt tttgagcaag gatgtttggc ggggccaaga 17040
ccagcaataa cacaggctgg ggcctgcgct tcccagcaa cgcgcgcgcc tggggcgcg 17100
agcgctccga ccaacaccca gtgcccgtgc ggcgggacta cgcgcgcgcc catcgacgcg 17160
acaaacgcgg cgcactggg cgcaccaccg cagtgacgc cagtgacgc agtgagcgcg 17220
aggcgcgcaa ctacacgccc acgcccgcac cagtgacgc agtgagcgc agtgagcgcg 17280
ccgtgggtgc cggagcccgc cgtatgcta aaatgaagag acggcgaggg gcgcgagcgc 17340
gtcggccacc cgcgcgccc ggcactgccc cccaacgcgc cggcgagcgc cgtgcttaac 17400
gcgacagtcg caccggccga cggcgggcca tgcggggccc tcgaaggctg gcccgggta 17460
ttgtcactgt gccccccagg tccaggcgac gagcgggccc cgcagcagcc ggggcccata 17520
gtgctatgac tcagggtcgc aggggcaacg tgtattgggt gcgcgactcg gttagcgcc 17580
tgccgctgccc cgtgcccacc cgcggggccc gcaactagat tgcaagaaaa aactacttag 17640

```

actcgtactg ttgtatgtat ccagcggcgg cggcgcgcaa cgaagctatg tccaagcgca 17700
aatcaaaaga agagatgctc caggatcatcg cgccggagat ctatggcccc ccgaagaagg 17760
aagagcagga ttacaagccc cgaaagctaa agcgggtcaa aaagaaaaag aaagatgatg 17820
atgatgaact tgacgacgag gtggaactgc tgcacgctac cgcgccagg cgacgggtac 17880
agtggaaaag tcgacgcgta aaacgtgttt tgcgacccgg caccaccgta gtctttacgc 17940
cgggtgagcg ctccacccgc acctacaagc gcgtgtatga tgaggtgtac ggcgacgagg 18000
acctgcttga gcaggccaac gagcgcctcg gggagtttgc ctacggaaaag cggcataagg 18060
acatgctggc gttgccgctg gacgagggca acccaacacc tagcctaaag cccgtaaacac 18120
tgcacgctat gctgcccgcg cttgcaccgt ccgaagaaaa gcgcggccta aagcgcgagt 18180
ctgggtgactt ggcacccacc gtgcagctga tggtagccca gcgccagcga ctggaagatg 18240
tcttgaaaaa aatgaccgtg gaacctgggc cggtagccga ggtccgcgtg cggccaatca 18300
agcaggtggc gccgggactg ggcgtgcaga tggagacaca aacgtccccg gttgcctcag 18360
gcaccagtat tgccaccgcc caggcggctc ctgcggccgc gtccaagacc tctacggagg 18420
cgggtggcga tgccgcggtg tttcgcgttt cagcccccg gcgcccgcg ggttcgagg 18480
tgcaaacgga cccgtggatg ctactgccc acctaccgcc cgtgctggcc cggccaatca 18540
agtacggcgc cgccagcgcg cctaccgccc aatagccct acatccttc attgcgccta 18600
cccccggtta tctggctac cctgggtggt ccagaagacg agcaactacc cgacgcgcaa 18660
ccaccactgg aacccgcgcg ggcgtgcgaa cctgctggcc cgtgctggcc cggccaatca 18720
tgccgagggg ggtcgcgaa ggaggcagga ttcttgctga tatggccctc acctgccgc 18780
ccagcatcgt ttaaaagccg gtctttgtgg gaatgcaccg taggaggggc atggccgggc 18840
tccgtttccc ggtgccggga ttccgaggaa accaccggcg gcggcgcgcg cgcacccgtc 18900
acggcctgac gggcggcatg cgtcgtgcgc accaccggcg gcggcgcgcg cgcacccgtc 18960
gcattgcgcg cggtatcctg cccctcctta ttccactgat cgcgcggcg attggcgccg 19020
tgcccggaat tgcacccgtg gccttgacag cgcagagaca ctgattaaaa acaagttgca 19080
tgtgaaaaaa tcaaaataaa aagtctggac tctcacgctc gcttggtcct gtaactattt 19140
tgtagaatgg aagacatcaa ctttgctct ctggccccgc gacacggctc gcgccggtc 19200
atgggaaact ggcaagatat cggcaccagc aatatgagcg gtggcgccctt cagctggggc 19260
tcgctgtgga gcggcattaa aaatttcggt tccaccgtta agaactatgg cagcaaggcc 19320
tggaacagca gcacaggcca gatgctgagg gataagttga aagagcaaaa tttccaacaa 19380
aaggtggtag atggcctggc ctctggcatt agcggggtgg cggacctggc ccgtagagga caaccaggca 19440
gtgcaaaata agattaacag taagcttgat ccccgccctc ccgtagagga gcctccaccg 19500
gcccgtggaga cagtgtctcc agaggggctg ggcgaaaagc gtccgcgccc cagcagggaa 19560
gaaactctgg tgacgcaaat agacgagcct ccctcgtagc aggagggcact aaagcaaggc 19620
ctgcccacca cccgtcccat cgcgcccatg gctaccggag tgctgggcca gcacacacc 19680
gtaacgctgg acctgcctcc cccgcgcgac acccagcaga aacctgtgct gccaggcccc 19740
accgccgttg ttgtaaccgg tccatagccgc gcgtccctgc gccgcgccc cagcggctcg 19800
cgatcgttgc gggccgtagc cagtggcaac tggcaaacag cactgaacag catcggtgg 19860
ctgggggtgc aatccctgaa ggcgcgacga tgcttctgaa tagctaacgt gtcgtatgtg 19920
tgcatgtat gcgtccatgt cgcgcgacga ggagctgctg agccgcgcg catgcacatc tcgggccagg 19980
ccaagatggc tacccttctg atgatgccgc agtggtctta catgcccacc gagacgtact 20040
acgcctcgga gtacctgagc cccgggctgg tgcagtttgc ccgcgccacc gagacgtact 20100
tcagcctgaa taacaagttt agaaacccca cgggtggcgcc tacgcacgac gtgaccacag 20160
accggtccca gcgtttgacg ctgcggttca tccctgtgga ccgtaggatg actgogtact 20220
cgtacaaggc gcggttcacc ctactgtgtg gtgataaccg tgtgctggac atggcttcca 20280
cgtactttga catccgcggc gtgctggaca ggggccctac ttttaagccc tactctggca 20340
ctgacctaaa cgccctggct cccaagggtg ccccaaatcc ttgcgaatgg gatgaagctg 20400
ctactgctct tgaaataaac ctagaagaag aggcagatga caacgaagac gaagtagacg 20460
agcaagctga gcagcaaaaa actcacgtat ttgggcaggc gccttattct ggtataaata 20520
ttacaaagga gggatttcaa ataggtgtcg aagggtcaa acctaaatat gccgataaaa 20580
catttcaacc tgaacctcaa ataggagaat ctcagtggtg cgaaactgaa attaatcatg 20640
cagctgggag agtccttaaa aagactaccc caatgaaacc atgttacggt tcatatgcaa 20700
aaccacaaa tgaaaatgga gggcaaggca ttcttgtaaa gcaacaaaaa ggaaagctag 20760
aaagtcaagt ggaaatgcaa tttttctcaa ctactgaggc gaccgcaggc aatggtgata 20820
acttgactcc taaagtggta ttgtacagt aagatgtaga tatagaaacc ccagacactc 20880
atatttctta catgcccact attaaggaag gtaactcacg ttagggacaa ttttattgg 20940
ctatgcccac caggcctaag ggtgttctgg cgggccaagc ataccagct aggtctgtga cagctatgat ccagatgtta 21000
acaacagcac gggtaatatg acagagcttt atgtggaatc aacttccaaa ttactgcttt cactggggag 21060
tagatttgca agacagaaac actgaagatg taaaacctaa aacaggtcag gaaaatggat 21120
atagaaccag gtacttttct cttaccaagg ttttcagata aaaatgaaat aagagttgga aataattttg 21180
gtgtgattaa tacagagact ttttcagata

```

ccatggaaat caatctaaat gccaacctgt ggagaaatctt cctgtactcc aacatagcgc 21420
tgtattttgcc cgacaagcta aagtacagtc cttccaacgt aaaaatttct gataacccaa 21480
acacctacga ctacatgaac aagcgagtggt tggctcccgg gttagtggac tgctacatta 21540
accttgagac acgctgggtcc cttgactata tggacaacgt caaccatttt aaccaccacc 21600
gcaatgctgg cctgcgctac cgctcaatgt tgctgggcaa tggctcgctat gtgccccttc 21660
acatccaggt gcctcagaag ttcttttgcca ttaaaaaacct ccttctcctg ccgggctcat 21720
acacctacga gtggaacttc aggaaggatg ttaacatggg tctgcagagc tccctaggaa 21780
atgacctaaag gggttgacgga gccagcatta agtttgatag catttgcttt tacgccacct 21840
tcttccccat ggcccacaac accgcctcca cgcttgaggc catgcttaga aacgacacca 21900
acgaccagtc ctttaacgac tatctctccg ccgccaacct gctctaccct ataccgcga 21960
acgctaccaa cgtgcccata tccatccctt cccgcaactg ggcggttttc cgcggtggg 22020
ccttcacgog ccttaagact aaggaaacct catcactggg ctcgggctac gacctttatt 22080
acacctactc tggctctata ccctacctag atggaacctt ttacctcaac cacaccttta 22140
agaaggtggc cattaccttt gactctcttg tcagctggcc ttgcaatgac cgctgctta 22200
cccccaacga gtttgaattt aagcgctcag ttgacgggga ggggttacaac gttgcccagt 22260
gtaacatgac caaagactgg ttcttggtac aaatgctagc taactacaac attggctacc 22320
agggtttcta tatccagag agctacaagg accgcatgta ctccttcttt agaaacttcc 22380
agcccatgag ccgtcaggtg gtggatgata ctaaatataa ggactacca atgctgggca 22440
tcctacacca acacaacaac tctggatttg ttggctacct tgccccacc atgctgcaag 22500
gacagcccta cctgctaac ttccctatc cgcttatagg caagaccgca gttgacagca 22560
ttaccagaa aaagtttctt tgcgatcgca coctttggcg catcccatc tccagtaact 22620
ttatgtccat gggcgactc acagacctgg gccaaaacct gctctacgac aactccgccc 22680
acgctagtag catgactttt gaggtggatc ccatggacga gccaccctt ctttatgttt 22740
tggttgaagt ctttgacgtg gtccgtgtgc accggccgca ccgctggctc atcgaaaccg 22800
tgtacctgag caccgcccctc tcggcgggca acgccaacac acgcaacacg aagcaacac 22860
aacaacagct ccgcccagtg gctccagtga gcaggaaactg aaagccattg tcaaagatct 22920
tggttggtggg ccatattttt tgggcaacta tgacaagcgc tttccaggct ttgtttctcc 22980
acacaagctc gctgcgcca tagtcaatac ggccggtcgc gagactgggg gcgtacactg 23040
gatggccttt gcctggaacc cgactcaaa aactgtctac ctctttgagc cctttggctt 23100
ttctgaccag cgactcaagc aggtttacca gtttgagtac gagtcaactc tgcgcccgtg 23160
cgccattgct tcttcccccg accgctgtat aacgctggaa aagtccacct aaagcgtaca 23220
ggggcccaac tcggcgccct gtggactatt ctgctgcag tttctccacg ctttgccaa 23280
ctggcccaaa actcccagtg atcacaacct caccatgaac cttattaccg gggtaacca 23340
ctccgctc cactcaagc cgccctactt ccgcagccac agtgcgcaga aacagctcta 23400
cagcttctctg gagcgccact cgccctactt cccgagccac agtgccgaga ttaggagcgc 23460
cacttctttt ttttatttgg aactcaatgt aaaataatgt accccaccct ttgcccgtctg 23520
aggcaaatgc ttttatttgg ggttctgccc cgcatcgcta tgcgccactg gtaggagac 23580
cgccgtttaa aaatcaagg tgcctccact aaactcaggc acaaccatcc gcgggagctc 23640
ggtgagatac tgggtgttag tgcctccact catcaccaac gcgttttagc gctggggcgc 23700
ggtgaagttt tcaactccaca ggctgcgcac gccctgcgag cgcgagttgc gatacacagg 23760
cgatatcttg aagtgcagtg tggggcctcc gtgggtgcag ctggccagca cgctcttgct 23820
gttgagcac tggaacacta tcagcgccgg gttgctcagg gctgactcgc tcaactttgg 23880
ggagatcaga tccgctcca ggtcctccgc aggtctttgag ttgactcgc accgtagtgg 23940
tagctgctt cccaaaaagg gcgctgccc gttaggatac agcgcctgca taaaagcctt 24000
catcaaaaagg tgaccgtgcc cggtctgggc gcttccagag aagaacatgc cgcaagactt 24060
gatctgctta aaagccacct gagcctttgc gctctgcag gttcttcag atcttggcct 24120
gccgaaaaac tgattggcgc gacaggccgc gtctgtcacg atttcaatca cgtcggtgtt 24180
ggagatctgc accacatttc ggccccaccg gttcttcag atcttggcct tgctagactg 24240
ctccttcagc gcgctgccc cggtttcgct cgctcacatc atttcaatca cgtgctcctt 24300
atctatcata atgcttccgt gtgacactt aagctcgct tccatctcag cgcagcgtg 24360
cagccacaac gcgagcccg tgggctcgtg atgctttag gtcacctctg caaacgactg 24420
caggtacgccc tgcaggaatc gcccatcat cgtcacaaag gtcctgttgc tgggtgaaggt 24480
cagctgcaac ccgctgctc cctcgctcag ccaggtcttg cctacggccg ccagagcttc 24540
cacttggtca ggcagtagtt tgaagtccgc ctttagatcg ttatccaggt ggtacttctc 24600
catcagcgc gcgagcctt ccatgccctt cctccacgca gacacgatcg gcacactcag 24660
cgggttcctc accgtaattt cactttccgc ttcgctgggc tcttctctt cctcttgcgt 24720
cgcatacca cgcgccactg ggtcgtcttc attcagccgc cgactgtgc gcttacctc 24780
tttgccatgc ttgattagca ccggtgggtt gctgaaacct accatttgta gctccacatc 24840
ttctctttct tcctcgctgt ccacgattac ctctgtgat ggcggggcgt cgggcttggg 24900
agaaggcgc ttctttttct tcttgggcgc aatggccaaa tccgcccgcg aggtcgatgg 24960
ccgcccggctg ggtgtgcgc gcaccagcgc gtcttctgat gactcttct cgtcctcgga 25020
ctcgatacgc cgcctcatcc gcttttttgg gggcgcccgg ggaggcggcg gcgacgggga 25080

cggggacgac acgtcctcca tgggtggggg acgtcgcgcg gcaccgcgct cgcgctcggg 25140
ggtggtttcg cgctgctcct cttcccgaact ggccatttcc ttctcctata ggcagaaaaa 25200
gatcatggag tcagtcgaga agaaggacag cctaaccgcc ccctctgagt tcgccaccac 25260
cgctccacc gatgcccca acgcgcctac caccttcccc gtcgaggcac ccccgcttga 25320
ggaggaggaa gtgattatcg agcaggaccc aggttttcta agcgaagacg acgaggaccg 25380
ctcagtagca acagaggata aaaagcaaga ccaggacaac gcagaggcaa acgaggaaaca 25440
agtcggggcg ggggacgaaa ggcattggca ctacctagat gtgggagacg acgtgctgtt 25500
gaagcatctg cagcgccagt gcgccattat ctgcgacgcg ttgcaagagc gcagcgatgt 25560
gccctcgcc atagcggatg tcagccttgc ctacgaacgc cacctattct caccgcgcgt 25620
acccccaaa cgccaagaaa acggcacatg cgagcccaac ccgcgcctca acttctaccc 25680
cgtatttggc gtgccagagg tgcttgccac accgcagccg agcggacaag cagctggcct tgcggcaggg 25740
accctatcc tgccgtgcca accgcagccg cgaagtgcc aaaaatcttg agggctcttg 25800
cgctgtcata cctgatatcg cctcgctcaa cgaagtgcc aaaaatcttg agggctcttg 25860
acgcgacgag aagcgcgcgg caaacgctct gcaacaggaa aacagcgaaa atgaaagtca 25920
ctctggagtg ttggtggaac tcgagggtga caacgcgcgc ctagccgtac taaaacgcag 25980
catcgaggtc acccactttg cctaccgggc acttaacctt ccccccaagg tcatgagcac 26040
agtcattgag gagctgatcg tgcgcgctgc gcagccctg gagagggatg caaatttgca 26100
agaacaaaca gaggaggggc taccgcagat tggcgacgag cagctagcgc gctggcttca 26160
aacgcgcgag cctgcccact tggaggagcg acgcaaaact atgatggcg cagtgtctgt 26220
taccgtggag cttgagtgca tgcagcggtt ctttgctgac ccggagatgc agcgcaagct 26280
agaggaaaca ttgcactaca cctttcgaca gggctacgta cgccaggcct gcaagatctc 26340
caacgtggag ctctgcaacc tggtctccta ccttggattt cgcgactacg tccgcgactg 26400
gcaaaacgtg cttcattcca cgctcaaggg cgaggcgcg cgcgactacg tccgcgactg 26460
cgtttactta tttctatgct acactggca gacggccatg ggcgtttggc agcagtgtct 26520
ggaggagtgc aacctcaagg agctgcagaa actgctaaag caaaacttga aggaactatg 26580
gacggccttc aacgagcgct ccgtggcgcg gcacctggcg gacatcattt tccccgaacg 26640
cctgtctaaa accctgcaac agggctctgcc agacttcacc agtcaaaagca tgttgcaaga 26700
ctttaggaac tttatcctag agcgctcagg aatcttgccc gccacctgct gtgcacttcc 26760
tagcgacttt gtgcccatta agtaccgcca atgcccctcg ccgctttggg gccactgcta 26820
ccttctgcag ctgcccactt accttgctca cactctgac caactatgc accccgcacc gctccctggt 26880
tgacgggtcta ctggagtgtc actgtcgtg caactatgc accccgcacc gctccctggt 26940
ttgcaattcg cagctgctta acgaaagtca aattatcggt acctttgagc tgcagggtcc 27000
ctcgccctgac gaaaagtccg cggctccggg gttgaaactc actccggggc tgtggagctc 27060
ggcttacctt cgcaaatttg tacctgagga ctaccacgcc cagcagatta ggttctacga 27120
agaccaatcc cgcccgccaa atgcggagct taccgcctgc gtcatatccc agggccacat 27180
tcttggccaa ttgcaagcca tcaacaaagc ccgccaagag tttctgctac gaaagggagc 27240
gggggtttac ttggacccc agtccggcga ggagctcaac ccaatcccc cgccgcccga 27300
gccctatcag cagcagccgc gggcccttgc ttcccaggat ggcacccaaa aagaagctgc 27360
agctgccgcc gccacccacg gacgaggagg aatactggga cagtcaggca gaggaggttt 27420
tgagcagagga ggaggaggac atgatggaag actgggagag cctagacgag gaagcttccg 27480
aggtcgaaga ggtgtcagac gaaacaccgt caccctcggt cgcattcccc tcgcccggcg 27540
cccagaaatc ggcaaccggt tccagcatgg ctacaacctc cgctcctcag gcgcgcggcg 27600
cactgcccgt tcgcccagcc aaccgtagat gggacaccac tggaaaccagg ccggttaagt 27660
ccaagcagcc gccgcccgtta gcccaagagc aacaacagcg ccaaggctac cgtcatggc 27720
gcgggcacaa gaacgccata gttgcttgc tgcaagactg tgggggcaac atctccttcg 27780
cccgcgctt tcttctctac catcacggcg tggccttccc ccgtaacatc ctgcattact 27840
accgtcatct ctacagccca tactgcaccg gcggcagcgg cagcggcagc aacagcagcg 27900
gccacacaga agcaaaaggc accggatagc aagactctga caaagcccaa gaaatccaca 27960
gcggcgcgag cagcaggagg aggagcgctg cgtctggcg ccaacgaacc cgtatcgacc 28020
cgcgagctta gaaacaggat ttttcccact ctgtatgcta tatttcaaca gagcagggcg 28080
caagaacaag agctgaaaat aaaaaacagg tctctcgat ceetcaccog cagctgctgt 28140
tatcacaaaa gcgaagatca gcttcggcgc acgctggaag acgcgaggcg tctcttcagt 28200
aaatactgag cgctgactct taaggactag tttcgcgcc tttctcaaat ttaagcgcg 28260
aaactacgtc atctcagagc ggcagagcgc ggcagagcgc ctgtctcag cccattatg 28320

ttcacgcctc gtcaggcaat cctaactctg cagacctcgt cctctgagcc gcgctctgga 28860
ggcattggaa ctctgcaatt tattgaggag tttgtgcat cggctctactt taacccttc 28920
tcgggacctc ccggccacta tccggatcaa tttattccta actttgacgc ggtaaaggac 28980
tcggcgagcg gctacgactg aatgttaagt ggagaggcag agcaactgcg cctgaaacac 29040
ctgggtccact gtcgcccga caagtgttt gcccgcgact ccggtgagtt ttgctacttt 29100
gaattgcccc gtagcctgat tcgggagttt ggcgacggcg tccggcttac cgcccaggga 29160
gagcttgccc gtgttctcac tgtgatttgc aactgtccta accttgagtt tgagcgggac 29220
aggggacctt gtcacatcct cttacacttt ttcatacatt gcccaagaat aaagaatcgt 29280
ttaattaatt gccacatcct cttacacttt aattgcagaa aatttcaagt catttttcat 29340
ttgtgttatg tttcaacgtg tttatttttc tatacagatc accgtacctt aatcaaacctc 29400
tcagtagtat agccccacca ccacatagct cctcccaaca cacagagtac acagtccttt 29460
acagaacctt agtattcaac ctgccacctc cctcccaaca cacagagtac acagtccttt 29520
ctccccgggt ggccttaaaa agcatcatat catgggtaac gatattaata aactccccgg 29580
tattccacac ggtttcctgt cgagccaaac gctcatcagt gatattaata aactccccgg 29640
gcagctcact taagtctcatg tcgctgtcca gctgctgagc cacaggctgc tgcctcaactt 29700
gcgggttgctt aacgggaggc gaaggagaag tccacgccta catgggggta gaggcctgt 29760
cgtgcatcag gatagggcgg tgggtgtgca gcagcgcgcg aataaaactgc tgccgcggcc 29820
gctcgcctct gcaggaatac aacatggcag tgggtcctc agcgatgatt cgcaccggcc 29880
gcagcataag gcgcttctg ctcggggcac agcagcgac cctgatctca cttaaatcag 29940
cacagtaact gcagcacagc accacaatat tgttcaaaa cccacagtgc aaggcctgt 30000
atccaaagct catggcgggg accacagaac ccacgtggcc atcataccac aagcgcagg 30060
agattaagtgc gcgacctc caccacacgc tggacataaa cattacctct tttggcatgt 30120
tgtaattcac cactcccg taccatataa acctctgatt aaacatggcg ccatccacca 30180
ccatcctaaa ccagctggcc aaaacctgcc cgccggctat acactgcagg gaaccgggac 30240
tggaacaatg acagtggaga gccaggaact gcatacact cctcaggatt aaccccacac 30300
tatcaatgtt ggcacaacac aggcacacgt gccacacact cctcaggatt aatccacac 30360
cccgcgtagt aaccatatcc cagggaacaa cccattcctg aatcagcgta aatccacac 30420
tgcagggaag acctcgacg taactcacgt tgtgcttctg ctcaaaggga ggtagacgat 30480
gcagcggatg acctccag atggtagcgc cgagacaacc gagatcgtgt gagtagacgat 30540
ccctactgta cggagtgcgc atatttctg aagcaaaacc aggtgcgggc gtcagccaa 30600
atggaacgcc ggacgtagtc ccgcttagat cgctctgtgt agtagttgta gtatattccac 30660
gatctgcgtc tccggtctcg ccccttggt tgggttctta tgtaaaactcc ttcagcgcc 30720
tctctcaaag catccaggcg ccccgagaa taagccacac ccagccaacc tacacattcg 30780
gctgccctga taacatccac caccgcagaa agagctggaa gaacctggt tttttttta 30840
ttctgcgagt cacacacggg agggagcggga aagatgtatt aagtgaaacgc gctccctcc 30900
ttccaaaaga tttccaaa cagccaaaga acagataatg gcatttgtaa gatgttgac 31020
gggtggcggtg tcaaaactca cggccctcac gtccaagtgg acgtaaaggc taacccttc 31080
aatggcttcc tctctataa acattccagc acctcaacc atgcccacac aattctcatc 31140
agggtgaatc tctcaatat ctctaagcaa atcccgaaata ttaagtccgg ccattgtaaa 31200
tcgccacact agagcgccct ccaccttcag cctcaagcag cgaatcatga ttgcaaaaa 31260
aatctgctcc cacagacctg tataagattc aaagcggaa cattaacaaa aataccgga 31320
tcaggttctt ccccttcgag ggcagctga aagaaacca cactgattat gaggcagat 31380
tcccgtaggt ccccttcgag aaccttgaca cgtagccccg atgtaagctt tgttgcatgg gcggcgat 31440
gcggccactt ccccgccagg aaccttgaca cgtagccccg atgtaagctt tgttgcatgg gcggcgat 31500
ctcgagcta tgctaaccag cgtagccccg atgtaagctt tgttgcatgg gcggcgat 31560
aaaatgcaag gtgctgctca aaaaatcagg caagcctcg accaccacag aaaaagacac 31620
gtagtcatgc tcatgcagat aaaggcagg ctgcataaac acaaaataaa ataacaaaa 31680
catttttctc tcaaacatgt ctgcttaca acaggaaaaa caacccttat aagcataaga 31740
aacattttaa cattagaagc ctgtcttaca caccgtgatt aaaaagcacc 31800
cggactacgg ccatgccggc gtgacgtaa ataagtgaag actcggtaaa caccatcagg 31860
accgacagct cctcggtcat gtcaggatc gaaatagccc gggggaatac ataccgcag 31920
tgattcatcg gtcagtgcta aacattacag cccccatagg aggtataaca aaattaatag gagagaaaa 31980
gcgtagagac aacattacag cccccatagg aggtataaca aaattaatag gagagaaaa 32040
cacataaaca cctgaaaaac cctcctgcct aggcataaac gaccctccc gctccagaa 32100
aacataacag gcttcacagc ggcagcctaa cagtcagcct taccagtaaa aaagaaaa 32160
tattaaaaaa acaccactcg acacggcacc agctcaatca gtcacagtgt aaaaaagg 32220
caagtgcaga gcgagtatat ataggactaa aaaatgacgt aacgggttaa gtccacaaa 32280
aacacccaga aaaccgcag cgaacctacg cgaagcaaac aaagccaaa aaccacaa 32340
ttctcaaat cgtcacttcc gttttccac gttacgtaac ttccatttt cccgttcc 32400
caattcccaa cacatacaag ttactccgac ctaaaacctc cgtcaccgcg aatccaaa 32460
cgccccggc caggtcacaa actccacccc ctcattatca tattggcttc 32480
aaggtatatt attgatgatg

<210> 24
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 24
ctcaacaatt gtggatccgt actcc

25

<210> 25
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 25
gtgctcagca gatcttgcca ctgtg

25

<210> 26
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 26
ggcgcgttcg gatccactct ctcc

25

<210> 27
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 27
ctacatgcta ggcagatctc gttcggag

28

<210> 28
<211> 1240
<212> DNA
<213> adenovirus

<400> 28
ggatccactc tcttccgcat cgctgtctgc gagggccagc tgttgggggtg agtactccct 60
ctgaaaagcg ggcattgactt ctgcgctaag attgtcagtt tccaaaaaacg aggaggattt 120
gatattcacc tggcccgcggt tgatgccttt gaggggtggcc gcatccatct ggtcagaaaa 180
gacaatcttt ttgttgtcaa gcttggtggc aaacgacccg tagaggcggt tggacagcaa 240
cttggcgatg gagcgacagg tttgggtttt gtcgcgatcg gcgcgctcct tggccgcgat 300
gttttagctgc acgtattcgc gcgcaacgca ccgccattcg ggaaagacgg tgggtgcgctc 360
gtcgggcacc aggtgcacgc gccaaacggc gttgtgcagg gtgacaaggc caacgctggt 420
ggctacctct ccgcgtaggc gctcgttggc ccagcagagg cggccgacct tgcgcgagca 480
gaatggcggt agggggtcta gctgcgtctc gtccggggggg tctgcgtcca cggtaaagac 540
cccgggcagc aggcgcgcgt cgaagtagtc tatcttgcatt ccttgcaagt ctacgcgctg 600

```

ctgccatgcg cgggcgggcaa ggcgcgcgtc gtatgggttg agtgggggac cccatggcat 660
gggggtgggtg agcgcgaggg cgtacatgcc gcaaagtgcg taaacgtaga ggggctctct 720
gagtattcca agatatgtag ggtagcatct tccaccgagg atgctggcgc gcacgtaac 780
gtatagttcg tgcgaggag cgaggaggtc gggaccgagg ttgctacggg cgggctgctc 840
tgctcggaag actatctgcc tgaagatggc atgtgagttg gatgatatgg ttggacgctg 900
gaagacgttg aagctggcgt ctgtgagacc taccgctca cgcacgaagg aggcgttaga 960
gtcgcgcagc ttgttgacca gctcggcggg gacctgcacg tctagggcgc agtagtccag 1020
ggtttcttg atgatgtcat acctatctcg tccctttttt tccacagct cgcggttag 1080
gacaaactct tcgcggtctt tccagtactc ttggatcgga aaccgcgcg cctccgaacg 1140
agatccgtac tccgccgccg agggacctga gcgagtcgcg atcgaccgga tcggaaaacc 1200
tctcgagaaa ggcgtctaac cagtcacagt cgcaagatct 1240

```

<210> 29

<211> 8383

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid pDV60

<400> 29

```

gacggatcgg gagatctccc gatccccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gtagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgtttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaataca ttacggggtc attagttcat agcccatata 300
tggaagtccg cgttacataa cttacggtaa atggcccgcg tggctgaccg cccaacgacc 360
cccgcccat t gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggg aaactgcccc cttggcagta catcaagtgt 480
atcatatgcc aagtagcgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catgggtgatg cggtttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccaccccc ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactcgcg cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatcaga ctactatag ggagacccaa gcttgggtacc 900
gagctcggtt ccaactctct cgcacgtcgt gtctgcgagg gccagctggt ggggtgagta 960
ctccctctga aaagcgggca tgacttctgc gctaagattg tcagtttcca aaaacgagga 1020
ggatttgata ttacacctggc ccgcggtgat gcctttgagg gtggccgcat ccatctggtc 1080
agaaaagaca atctttttgt tgtcaagctt ggtggcaaac gacctgtaga gggcggttga 1140
cagcaacttg gcgatggagc gcagggtttg gtttttgcg cgatcggcgc gctccttggc 1200
cgcgatgttt agctgcacgt attcgcgcgc aacgcaccgc cattcgggaa agacgggtgt 1260
gcgctcgctg ggcaccaggt gcacgcgcca accgcggtg tgcagggtga caaggtcaac 1320
gctgggtggc acctctcgcg gtaggcgctc gttgggtccag cagaggcggc cgcccttgcg 1380
cgagcagaat ggcggtaggg ggtctagctg cgtctcgtcc ggggggtctg cgtccacggt 1440
aaagaccccg ggcagcaggc gcgcgtcgaa gtagtctatc ttgcatcctt gcaagtctag 1500
cgctctgctc catgcgcggg cggcaagcgc gcgctcgtat ggggttagtg ggggacccca 1560
tgccatgggg tgggtgagcg cggaggcgta catgccgcaa atgtcgtaaa cgtagagggg 1620
ctctctgagt attccaagat atgtagggtg gcatcttcca ccgcggatgc tggcgcgcac 1680
gtaatcgtat agttcgtgcg agggagcgag gaggtcgga ccgaggttgc tacgggcggg 1740
ctgctctgct cggaagacta tctgcctgaa gatggcatgt gagttggatg atatgggttg 1800
acgttggaag acgttgaagc tggcgtctgt gagacctacc gcgtcacgca cgaaggaggc 1860
gtaggagtcg cgcagcttgt tgaccagctc ggcggtgacc tgcacgtcta gggcgagta 1920
gtccagggtt tccttgatga tgtcatactt atcctgtccc ttttttttcc acagctcgcg 1980
gttgaggaca aactcttcgc ggtctttcca gtactcttg atcggaacc cgtcggcctc 2040
cgaacgagat ccgtactccg ccgcgcgagg acctgagcga gtccgcacgt accggtatcg 2100
aaaaactctc gagaaaggcg tctaaccagt cacagtcgca agatccaaga tgaagcgcg 2160
aagaccgtct gaagatacct tcaaccccg gtatccatat gacacggaaa ccggtcctcc 2220
aactgtgcct tttcttactc ctccctttgt atcccccaat gggtttcaag agagtcccc 2280
tggggtactc tctttgcgcc tatccgaacc tctagttacc tccaatggca tgcttgcgct 2340
caaatgggc aacggcctct ctctggacga ggccggcaac cttacctccc aaaatgtaac 2400
cactgtgagc ccacctctca aaaaaaccaa gtcaaacata aacctggaaa tatctgcacc 2460

```



```

cctcacagtt acctcagaag ccctaactgt ggctgcccgc gcacctctaa tggctgcggg 2520
caacacactc accatgcaat cacaggcccc gctaaccgtg cactgactcca aacttagcat 2580
tgccacccaa ggacccctca cagtgtcaga aggaaagcta gccctgcaaa catcaggccc 2640
cctcaccacc accgatagca gtacccttac tatcactgcc tcacccctc taactactgc 2700
cactggtagc ttgggcattg acttgaaaga gccattttat acacaaaatg gaaaactagg 2760
actaaagtac ggggctcctt tgcattgtaac agacgacctc aacactttga ccgtagcaac 2820
tggctccaggt gtgactatta ataatacttc cttgcaaaact aaagttactg gagccttggg 2880
ttttgattca caaggcaata tgcaacttaa tgtagcagga ggactaagga ttgattctca 2940
aaacagacgc cttatacttg atgttagtta tccgtttgat gctcaaaacc aactaaatct 3000
aagactagga cagggccctc tttttataaa ctcagcccac aacttgagga ttaactacaa 3060
caaaggcctt tacttgttta cagcttcaaa caattccaaa aagcttgagg ttaacctaa 3120
cactgccaag gggttgatgt ttgacgtac agccatagcc attaatgcag gagatgggct 3180
tgaatttggg tcacctaatg caccaaacc aaatccctc aaaacaaaaa ttggccatgg 3240
cctagaattt gattcaaaaca aggctatggt tcctaaacta ggaactggcc ttagtttga 3300
cagcacaggt gccattacag taggaaacaa aataaatgat gatgctaaac tctacttggg 3360
accagctcca tctcctaact gtagactaaa tgcagagaaa gttttggctg ttaaggcag 3420
cttaacaaaa tgtggcagtc aaatacttgc tacagtttca gttttggctg tttaaaggcag 3480
tttggctcca atatctggaa cagttcaaaag tgctcatctt attataagat ttgacgaaaa 3540
tggagtgtca ctaaacaaat ccttctgga ccagaatat tggaaactta gaaatggaga 3600
tcttactgaa ggccacagct atacaaacgc tgttggattt atgctaacc tatcagctta 3660
tccaaaatct cagggtaaaa ctgcaaaaag taacattgtc agtcaagttt acttaaacgg 3720
agacaaaaat aaacctgtaa cactaaccat ctatgtcatt ggtacacagg aaacaggaga 3780
cacaactcca agtgcatact cctcttacac ttcatgggac tggctctggcc acaactacat 3840
taatgaaata tttgccacat cttagaggcc ctattctata attgcccag aataaaagaa 3900
gcggccgctc gagcatgcat gtgccttota gttgccagcc atctgttgtt aatgctagag 3960
ctcgctgac agcctcgact gaaggtgcca ctcccactgt cctttcctaa tggccctccc 4020
ccgtgccttc cttgacctg agtaggtgtc attctattct ggggggtggg gtggggcagg 4080
aaattgcac gcattgtctg agtaggtgtc agtaggtgtc ggggggtggg gtggggcagg 4140
acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgag gtggggccta 4200
tggcttctga ggccgaaaga accagctggg gctctagggg gtatcccac gcgcctgtg 4260
gcggcgctt aagcgcgcg ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca 4320
gcggccctag gcccgctcct ttcgctttct tcccttcctt tctcgccag tctcgccag 4380
ttccccgtca agctctaaat cggggcatcc ctttaggtt ccgatttagt gctttacggc 4440
acctcgacct caaaaaactt gattagggtg atggttcacg tagtgggcca tgcgccctg 4500
agacggtttt tgcgcctttg acgttggagt cctatctcgg tctattcttt tgatttataa 4560
aaactggaac aactactcaac cctatctcgg tctattcttt tgatttataa aaaatttaac 4620
ggatttcggc ctattgggta aaaaatgagc tagtttaaca aaagtcccca ggctcccag 4680
tctgtggaat gtgtgtcagt taggggtgtg aaagtcccca ggctcccag 4740
gtatgcaaaag catgcatctc aattagtcag caaccagggtg tggaaagtcc ccaggctccc 4800
cagcaggcag aagtatgcaa agcatgcatc caaccagggtg tggaaagtcc ccaggctccc 4860
taactccgcc catcccgccc ctaactccgc ccagttccgc agcaaccata gtcccgcgcc 4920
gactaatttt ttttatttat gcagaggccg aggcgcgccc ccattctccg ccccatggct 4980
agtagtgagg aggttttttt ggaggcctag gcttttgcaa aaagctcccg ggagcttgta 5040
tatccatttt cggatctgat caagagacag gatgaggatc gtttcgcatg attgaacaag 5100
atggattgca cgcaggttct ccggccgctt ggggtggagag gctattcggc tatgactggg 5160
cacaacagac aatcggctgc tctgatgccc cgtgttccg gctgtcagcg caggggcgcc 5220
cggttctttt tgtcaagacc gacctgtccg gtgcccgtga tgaactgcag gacgaggcag 5280
cgcggctatc gtggctggcc acgacggggc ttccctgccc agctgtgtct gacgttgtca 5340
ctgaagcggg aagggactgg ctgctattgg gcgaagtgcc ggggcaggat ctcctgtcat 5400
ctcaccttgc tccgtccgag aaagtatcca tcatggctga tgcaatgcgg cggctgcata 5460
cgcttgatcc ggctacctgc ccattcgacc aggatgatct ggacgaagag gatcaggggc 5520
gtactcggat ggaagccggg cctgtcgatc aggcggcctc gcccgcaggc gaggatctcg 5580
tcgcgccagc cgaactgttc gccaggctca agtcggcctc gggaaatggc cgcttttctg 5640
tcgtgacca tggcgatgcc tgcttgccga atatcatggt ggaatggcgc cgttttctg 5700
gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttggcta 5760
cccgatgata tgctgaagag cttggcggcg aatgggctga ccgcttctc gtgctttacg 5820
gtatcgccgc tcccgattcg cagcgcatcg ccttctatcg ccttcttgac gagttcttct 5880
gagcgggact ctgggggtcg aaatgaccga ccaagcgacg cccaacctgc catcacgaga 5940
tttcgattcc accgcgcct tctatgaaag gttgggcttc ggaatcgctt tccgggacgc 6000
cggttgatg atcctccagc gcggggatct catgctggag ttcttcgccc accccaactt 6060
gtttattgca gcttataatg gttacaataa atcacaatt atcacaataa 6120
agcatttttt tcaactgcatt ctagtgtgtg tttgtccaaa ctcactcaatg tatcttatca 6180

```


tgtctgtata	ccgtcgacct	ctagctagag	cttggcgtaa	tcattggtcat	agctggttcc	6240
tgtgtgaaat	tgttatccgc	tcacaattcc	acacaacata	cgagccggaa	gcataaagt	6300
taaagcctgg	ggtgccta	gagtgagcta	actcacatta	attgcgttgc	gctcactgcc	6360
cgctttccag	tcgggaaacc	tgctgtgcca	gctgcattaa	tgaatcggcc	aacgcgcggg	6420
gagaggcgg	ttgcgtattg	ggcgctcttc	cgcttccctc	ctcactgact	cgctgcgctc	6480
ggtcggtcgg	ctgcggcgag	cggtatcagc	tcactcaaa	gcggttaata	ggttatccac	6540
agaatcagg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	6600
ccgtaaaaag	gccgcgttgc	tggtcggttt	ccataggctc	cgccccctg	acgagcatca	6660
caaaaatcga	cgctcaagtc	agagggtggc	aaacccgaca	ggactataaa	gataccaggc	6720
gtttccccct	ggaagctccc	tcgtgcgctc	tcctgttccg	accctgccc	ttaccggata	6780
cctgtccgcc	tttctccctt	cggaagcgt	ggcgctttct	caatgctcac	gctgtaggta	6840
tctcagttcg	gtgtaggctg	ttcgctccaa	gctgggctgt	gtgcacgaac	cccccttca	6900
ccccgaccgc	tgcccttat	ccggttaacta	tcgtcttgag	tcacacccgg	taagacacga	6960
cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	agagcgagg	atgtaggcgg	7020
tgctacagag	ttcttgaagt	ggtggcctaa	ctacggctac	actagaagga	cagtatttgg	7080
tatctgcgct	ctgctgaagc	cagttacctt	cggaataaga	gttggtagct	cttgatccgg	7140
caaaacaacc	accgctggta	gcggtgggtt	ttttgtttgc	aagcagcaga	ttacgcgcag	7200
aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	gggtctgacg	ctcagtgga	7260
cgaaaactca	cgtaagggga	ttttggctcat	gagattatca	aaaaggatct	tcacctagat	7320
cccttttaaat	taaaaatgaa	gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	7380
tgacagttac	caatgcttaa	tcagtggagc	acctatctca	gcgatctgtc	tatttcgttc	7440
atccatagtt	gcctgactcc	ccgtcgtgta	gataactacg	atacgggagg	gcttaccatc	7500
tggtcccag	gctgcaatga	ccgcgcgaga	cccacgctca	ccggctccag	atctatcagc	7560
aataaaccag	ccagccggaa	gggcccagcg	cagaagtgg	cctgcaactt	tatccgcctc	7620
catccagctc	attaattgtt	gccgggaagc	tagagtaagt	agttcgccag	ttaatagttt	7680
gcgcaactgt	gttgccattg	ctacaggcat	cgtgggtgca	cgctcgctcg	ttgggtatgg	7740
ttcattcagc	tcgggttccc	aacgatcaag	gcgagttaca	tgatccccca	tggtgtgcaa	7800
aaaagcgggt	agctccttcg	gtcctccgat	cgttgtcaga	agtaagttgg	ccgcagtggt	7860
atcactcatg	gttatggcag	cactgcataa	ttctcttact	gtcatgccat	ccgttaagatg	7920
cttttctgtg	actggtgagt	actcaaccaa	gtcattctga	gaatagtgtg	tgccggcgacc	7980
gagttgctct	tgcccggcgt	caataccggga	taataccgcg	ccacatagca	gaactttaaa	8040
agtgtctcatc	attggaaaac	gttcttcggg	gcgaaaactc	tcaaggatct	cttttacttt	8100
gagatccagt	tcgatgtaac	ccactcgtgc	acccaactga	tcttcagcat	cttttacttt	8160
caccagcggt	tctgggtgag	caaaaacagg	aaggcaaaat	gccgcaaaaa	agggaaataag	8220
ggcgacacgg	aaatgttgaa	tactcatact	cttccttttt	caatattatt	gaagcattta	8280
tcagggttat	tgtctcatga	gcggatacat	atttgaatgt	atttagaaaa	ataaacaat	8340
aggggttccg	cgcacatttc	cccgaagaat	gccacctgac	gtc		8383

<210> 30

<211> 7960

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid pDV67

<400> 30

gacggatcgg	gagatctccc	gatcccttat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttgtgtgtt	ggaggctcgt	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgtttttgc	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttattaa	tagtaatcaa	ttacggggtc	attagttcat	agccccatata	300
tggtgttccg	cgttacataa	cttacggtaa	atggcccgcc	tggttgaccg	cccaacgacc	360
cccgcccat	gacgtcaata	atgacgtatg	ttcccatagt	aacgcccaata	gggactttcc	420
attgacgtca	atgggtggac	tattttacgtt	aaactgccca	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	octattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggtttttgg	agtacatcaa	tggcgctgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaatgggag	780
gtaggcgtgt	acgggtggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagc	900

gtttaaactt aagcttggta ccgagctcgg atccactctc ttccgcacg ctgtctgcga 960
gggccagctg ttgggggtgag tactccctct gaaaagcggg catgacttct gcgctaagat 1020
tgtcagtttc caaaaacgag gaggatttga tattcacctg gcccgcggtg atgcctttga 1080
gggtggccgc atccatcttg tcagaaaaga caatcttttt gttgtcaagc ttggtggcaa 1140
acgaccgta gaggcggtg gacagcaact tggcgatgga gcgcagggtt tggttttgt 1200
cgcgatcggc gcgctccttg gccgcgatgt ttagctgcac gtattcgcg gcaacgcacc 1260
gccattcggg aaagacggtg gtgcgctcgt cgggcaccag gtgcacgcgc caaccgcggt 1320
tgtgcagggt gacaaggtca acgctggtgg ctacctctcc gcgtaggcgc tcggtgggtcc 1380
agcagaggcg gccgcccctg cgcgagcaga atggcggtag ggggtctagc tgcgtctcgt 1440
ccgggggggtc tgcgtccacg gtaaaagacc cgggcagcag gcgcgcgtcg aagtagtcta 1500
tcttgcatcc ttgcaagtct agcgctgct gccatgcgcg ggcggcaagc gcgcgctcgt 1560
atgggttgag tgggggaccc catggcatgg ggtgggtgag cgcggaggcg tacatgccgc 1620
aaatgtcgta aacgtagagg ggctctctga gtattccaag atatgtagg tagcatcttc 1680
caccgcggt gctggcgcg gcgtgctctg ctcggaagac tatctgcctg aagatggcat 1740
gaccgagggt gctacggcg ggacgctgga agacgttgaa gctggcgctc gtgagacct 1800
gtgagttgga tgatatgggt ggacgctgga agacgttgaa gctggcgctc gtgagacct 1860
ccgcgtcacg cacgaaggag gcgtaggagt cgcgcagctt gttgaccagc tcggcggtga 1920
cctgcacgtc tagggcgcg tagtccaggg tttccttgat gatgtcatac ttatcctgtc 1980
cctttttttt ccacagctcg cggttgagga caaactcttc gcggtctttc cagtactctt 2040
ggatcgga aa cccgtcgcc tccgaacgag atccgtactc gcgcgcgag ggacctgagc 2100
gagtcgcgat cgaccggatc gga aaacctc tcgagaaagg cgtctaacca gtcacagtgc 2160
caagatccaa gatgaagcgc gcaagaccgt ctgaagatac cttcaacccc gtgtatccat 2220
atgacacgga aaccggtcct ccaactgtgc cttttcttac tccctccctt gtatccccc 2280
atgggtttca agagagtccc cctggggtac tctctttgcg cctatccgaa cctctagtta 2340
cctccaatgg catgcttgcg ctcaaaatgg gcaacggcct ctctctggac gaggccggca 2400
accttacctc ccaaaatgta accactgtga gccacctct caaaaaaac aagtcaaa 2460
taaacctgga aatatctgca cccctcacag ttacctcaga agccctaact gtgggtgccc 2520
ccgcacctct aatggctcg gcgaacacac taacatgca atcacaggcc ccgctaaccg 2580
tgacagactc caaacttagc attgccaccc ccctcacca gcttgggcat caaggtatg 2640
tagccctgca aacatcaggc gccactggta gcttgggcat tgacttgaag gagccattt 2700
cctcaccccc tctaactact ggaactaaagt ggactaaagt tttgcatgta acagacgacc 2760
atacacaaaa tggaaaacta gacgtagca actggtccag gtgtgactat taataatact tccttgcaaa 2820
taaacacttt gaccgtagca ggttttgatt cacaaggcaa tatgcaactt aatgtagcag 2880
ctaaagttag gattgattct caaaacagac gccttatact tgatgttagt tatccgtttg 2940
gaggactaag gattgattct caaaacagac gccttatact tgatgttagt tatccgtttg 3000
atgctcaaaa ccaactaaat ctaagactag gacagggccc tctttttata aactcagccc 3060
acaacttgga tattaactac acaaaaggcc tttacttgtt tacagcttca aacaattcca 3120
aaaagcttga ggttaacctt agcactgcca aggggttgat gtttgacgct acagccatag 3180
ccattaatgc aggagatggg cttgaatttg gttcacctaa tgcacaaaac acaaatcccc 3240
tcaaaacaaa aattggccat ggcctagaat ttgattcaaa caaggctatg gttcctaaac 3300
taggaactgg ccttagtttt gacagcacag gtgccattac agtaggaaac aaaaaaatg 3360
ataagctaac tttgtggacc acaccagctc catctcctaa ctgtagacta aatgcagaga 3420
aagatgctaa actcactttg gtcttaacaa aatgtggcag tcaaatactt gctacagttt 3480
cagttttggc tgttaaaggc agtttggtc caatatctgg aacagttcaa agtgctcatc 3540
ttattataag atttgacgaa aatggagtgc tactaaacaa ttccttctctg gaccagaat 3600
attggaactt tagaaatgga gatcttacty aaggcacagc ctatacaaac gctgttggat 3660
ttatgcctaa cctatcagct tatccaaaat ctacggtaa aactgcaaaa agtaacattg 3720
tcagtcaagt ttacttaaac ggagacaaaa ctaaacctgt aacactaacc attactacta 3780
acggtacaca ggaaacagga gacacaactc caagtgcata ctctatgtca tttcatggg 3840
actggtctgg ccacaactac attaatgaaa tatttgccac atcctcttac actttttcat 3900
acattgccc aagaataaaag aagcggccgc tcgagtctag agggcccgtt taaacccgct 3960
gatcagcctc gactgtgcct tctagtggc agccatctgt tgtttgcccc tccccggtg 4020
cttcttgac cctggaagggt gccactccca ctgtcctttc ctaataaaat gaggaattg 4080
catcgcatg tctgagtagg tgtcattcta ttctggggg tggggtgggg caggacagca 4140
agggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt 4200
ctgaggcgga aagaaccagc tggggctcta ggggggtatcc ccacgcgccc tgtagcgcg 4260
cattaagcgc ggccgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc 4320
tagcgccgc tctttcgtc ttcttccctt cctttctgcg cacgttcgdc ggctttcccc 4380
gtcaagctct aaatcggggc atcccttttag ggttccgatt tagtgcttta cggcacctcg 4440
accccaaaaa acttgatttag ggtgatgggt cacgtagtgg gccatcgccc tgaatagcgg 4500
tttttcgccc tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg 4560
gaacaacact caaccctatc tcgggtctatt cttttgattt ataagggtt ttggggattt 4620

```

cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat taattctgtg 4680
gaatgtgtgt cagtttaggt gtggaaagtc cccaggtccc ccaggcaggc agaagtatgc 4740
aaagcatgca tctcaattag tcagcaacca ggtgtggaaa gtccccaggc tccccagcag 4800
gcagaagtat gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc 4860
cgccccatccc gcccctaact ccgcccagtt cgcgccattc tccgccccat ggctgactaa 4920
ttttttttat ttatgcagag gccgaggccg cctctgcctc tgagctattc cagaagtagt 4980
gaggaggcctt ttttgagggc ctaggccttt gcaaaaaagct cccgggagct tgtatatcca 5040
ttttcggatc tgatcagcac gtgttgacaa ttaatcatcg gcatagtata tcggcatagt 5100
ataatacgac aaggtgagga actaaaccaa ggccaagtgt accagtgcg tcccggtgct 5160
caccgcgcgc gacgtcgccg gagcggtcga gttctggacc gaccggctcg ggttctcccg 5220
ggacttcgtg gaggacgact tcgcccgtgt ggtccgggac gacgtgacct tgttcatcag 5280
cgccggtccag taccaggtgg tgccggacaa caccctggcc tgggtgtggtg tgcgcggcct 5340
ggacgagatc tacgccgagt ggtcggaggt cgtgtccacg aacttccggg acgctccgg 5400
gccggccatg accgagatcg gcgagcagcc gtggggggcg gagttcgccc tgcgcgacct 5460
ggccggcaac tgcgtgcact tcgtggccga ggagcaggac tgacacgtgc tacgagattt 5520
cgattccacc gccgccttct atgaaaggtt gggcttcgga atcgttttcc gggacgcgg 5580
ctggatgac ctccagcgcg gggatctcat gctggagttc ttcccacc ccaacttgtt 5640
tattgcagct tataatggtt acaataaaag caatagcatc acaaatttca caaataaagc 5700
atttttttca ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt 5760
ctgtataccg tcgacctcta gctagagctt ggcctaactc tggatcatagc tgtttcctgt 5820
gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa 5880
agcctggggt gcctaattag tgagctaact cacattaatt gcgttcgct cactgcccgc 5940
tttccagtcg ggaacacctgt cgttctccgc gcattaatga atcgccaac gcgcggggag 6000
aggcggtttg cgtattgggc tatcagctca ctcaaaggcg actgactcgc tgcgctcgg 6060
cgttcggctg cggcgagcgg atcaggggat aacgcaggaa agcaaaaggc gtaatacgg 6120
taaaaaggcc gcgttgctgg cgtttttcca taggtccgc cagcaaaagg ccgcaaacg 6180
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 6240
tccccctgga agctccctcg tgccgtctcc gcttctcaa ctgcccgtta ccgataacct 6300
gtccgccttt ctcccttcgg gaagcgtggc gctccaagct tcttctcagc ttaggtatct 6360
cagttcggtg gccttatccg gtaactatcg gcttctcagc caggaacccc ccgttcagcc 6420
cgaccgctgc ataggtcggtc gctccaagct tcttgagtc aaccgggtaa gacacgactt 6480
atcgccactg gcagcagcca ctggtaacag gattagcaga ggcaggtatg taggcggtgc 6540
tacagagttc ttgaagtggg ggctaacta cggctacact agaaggacag tatttggtat 6600
ctgcgctctg ctgaagccag ttacctcgg aaaaagagtt ggtagctctt gatccggcaa 6660
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta gtcgcaaaa 6720
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaaacg 6780
aaactcacgt taagggattt tttaatcaat ctaaagtata tatgagtaaa cttggtctga 6840
tttaaattaa aaatgaagtt ttgaggtacc atctgtctat tctgaggtat ttcgttcatc 6900
cagttaccaaa tgcttaatca gtgaggcacc tatctcagcg cgggaggggt taccatctgg 6960
catagttgcc tgactccccg tcgtgtagat acctacgata cgggaggggt taccatctgg 7020
ccccagtgct gcaatgatac cgcgagacc acgctcaccg gctccagatt tatcagcaat 7080
aaaccagcca gccggaaggg ccgagcgcag aagtggctct gcaactttat ccgcctccat 7140
ccagtctatt aattgttgcc gggaagctag agtaagtagt tgcgagttat cccctccat 7200
caacggttgt gccattgcta caggcatcgt ggtgtcacgc tcgccagttat atagtttgcg 7260
attcagctcc gggtcccaac gatcaaggcg agttacatga tctctggttg gtatggcttc 7320
agcggttagc tccttcggtc ctccgatcgt tgtcagaagt tccccatgt tgtgcaaaaa 7380
actcatggtt atggcagcac tgcataatc tcttactgtc aagttggccg cagtgttatc 7440
ttctgtgact ggtgagtact caaccaagtc attctgagaa atgccatccg taagatgctt 7500
ttgctcttgc cggcgctcaa tacgggataa taccgcgcca tagtgtatgc ggcgaccgag 7560
gctcatcatt ggaaaacggt cttcggggcg aaaactctca catagcagaa ctttaaaagt 7620
atccagttcg atgtaaccca ctggtgcacc caactgatct aggatcttac ttagttttag 7680
cagcgtttct ggggtgagcaa aaacaggaag gcaaaatgcc ttagttttag 7740
gacacggaaa tgttgaatac tcatactctt cctttttcaa gcaaaaaagg gaataagggc 7800
gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 7860
ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc tagaaaaata aacaaatagg 7920

```

<210> 31

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 31

atgggatcca agatgaagcg cgcaagaccg

30

<210> 32

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 32

cactatagcg gccgcattct cagtcattct

30

<210> 33

<211> 7989

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid pDV69

<400> 33

gacggatcgg	gagatctccc	gatccoctat	ggcgcactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttgtgtgtt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgttttgcg	ctgcttcgog	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttattaa	tagtaatcaa	ttacggggtc	attagttcat	agcccatata	300
tggagttccg	cgttacataa	cttacggtaa	atggcccggc	tggtgaccg	cccaacgacc	360
cccgcccat	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacggg	aaactgcccc	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacggcc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtea	600
tcgctattac	catgggtgatg	cggttttggc	agtagatcaa	tgggctgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaatgggcg	780
gtaggcggtg	acgggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagc	900
gtttaaactt	aagcttggtg	ccgagctcgg	atccactctc	ttccgcatcg	ctgtctgcga	960
ggggccagctg	ttgggggtgag	tactccctct	gaaaagcggg	catgacttct	gcgctaagat	1020
tgtcagtttc	caaaaacgag	gaggatttga	tattcacctg	gcccgcgggtg	atgcctttga	1080
gggtggggcgc	atccatctgg	tcagaaaaga	caatcttttt	gttgtcaagc	ttggtggcaa	1140
acgacccgta	gagggcggtg	gacagcaact	tggcgatgga	gcgcagggtt	tggtttttgt	1200
cgcgatcggc	gcgctccttg	gccgcgatgt	ttagctgcac	gtattcgcgc	gcaacgcacc	1260
gccattcggg	aaagacgggtg	gtgcgctcgt	cgggcaccag	gtgcacgcgc	caaccgcggg	1320
tgtgcagggt	gacaaggta	acgctgggtg	ctacctctcc	gcgtaggcgc	tcgttgggtc	1380
agcagaggcg	gccgcccctg	cgcgagcaga	atggcggtag	ggggtctagc	tgctctcgt	1440
ccgggggggtc	tgctccacg	gtaaaagacc	cgggcagcag	gcgcgcgtcg	aagtagtcta	1500
tcttgcatcc	ttgcaagtct	agcgctcgt	gccatgcgcg	ggcggcaagc	gcgcgctcgt	1560
atgggttgag	tggtgggacc	catggcatgg	gggtgggtgag	cgcgaggcg	tacatgccgc	1620
aaatgtcgta	aacgtagagg	ggctctctga	gtattccaag	atatgtaggg	tagcatcttc	1680
caccgcggat	gctggcgcg	acgtaatcgt	atagttcgtg	cgagggagcg	aggaggtcgg	1740
gaccgagggt	gctacggcg	ggctgctctg	ctcggaagac	tatctgcctg	aagatggcat	1800
gtgagttgga	tgatatggtt	ggacgctgga	agacgttgaa	gctggcgtct	gtgagacct	1860
ccgcgtcacg	cacgaaggag	gcgtaggagt	cgcgacgctt	ggtgaccagc	tcggcggtga	1920
cctgcacgtc	tagggcgag	tagtccaggg	tttccctgat	gatgtcatac	ttatcctgtc	1980
cctttttttt	ccacagctcg	cggttgagga	caaaactctc	gcggctcttc	cagtactctt	2040
ggatcgga	cccgtcggcc	tccgaacgag	atccgtactc	cgccgcgag	ggacctgagc	2100
gagtcgcgat	cgaccggatc	ggaaaacctc	tcgagaaagg	cgtctaacca	gtcacagtcg	2160

caagatccaa gatgaagcgc gcaagaccgt ctgaagatac ettcaacccc gtgtatccat 2220
atgacacgga aaccgggtcct ccaactgtgc cttttcttac tectcccttt gtatccccc 2280
atgggtttca agagagtcct cctgggttac tctctttgcg cctatccgaa cctctagtta 2340
cctccaatgg catgcttgcg ctcaaaatgg gcaacggcct ctctctggac gaggccggca 2400
accttacctc ccaaaatgta accactgtga gccacacctc caaaaaaacc aagtcacaca 2460
taaacctgga aatatctgca cccctcacag ttacctcaga agccctaact gtggctgccc 2520
ccgacacctc aatggtcgcg ggcaacacac tcacctgca atcacaggcc ccgctaaccg 2580
tgcacgactc caaacttagc attgccaccc aaggaccctc cacagtgtca gaaggaaagc 2640
tagccctgca aacatcaggc cccctcacca ccaccgatag cagtaccctt actatcactg 2700
cctcaccccc tctaactact gccactggta gcttgggcat tgacttgaaa gagcccat 2760
atacacaaaa tggaaaacta ggactaaagt acggggctcc tttgcatgta acagacgacc 2820
taaacacttt gaccgtagca actggtccag gtgtgactat taataact tctttgcaaa 2880
ctaaagttac tggagccttg ggttttgatt caaagggcaa tatgcaactt aatgtagcag 2940
gaggactaag gattgattct caaaacagac gccttatact tgatgttagt tatccgcttg 3000
atgctcaaaa ccaactaaat ctaagactag gacaggggcc tctttttata aactcagccc 3060
acaacttgga tattaactac aacaaaggcc ttacttgggt gtttgacgct acagccatag 3120
aaaagcttga ggttaacctc agcactgcca aggggttgat gtttgcacgt acagccatag 3180
ccattaatgc aggatgggg cttgaatttg gttcacctaa tgcacaaaac acaaatcccc 3240
tcaaaacaaa aattggccat ggccatgaat ttgattcaaa caaggctatg gttcctaaac 3300
taggaactgg ccttagtttt gacagcacag gtcgccattac agtaggaaac aaaaataatg 3360
ataagctaac tttgtggacc ggtccaaaac cagaagccaa ctgcataatt gaatacggga 3420
aacaaaaccc agatagcaaa ctaactttta tccttgtaaa aaatggagga attgttaag 3480
gatatgtaac gctaattgga gcctcagact acgttaacac cttattttaa aacaaaaatg 3540
tctccattaa tgtagaacta tactttgatg agcaaacccg cacttggta tatattacca gactcatctt 3600
ctcttaaaac agatctagaa ctaaaataca agcacttgat agcacttgat tgactttagt gcaagaggtt 3660
ttatgccaag tactacagcg tatccatttg tccttcctaa tgcgggaaca cataatgaaa 3720
attatatttt tggccaatgc tactacaaag caagcgatgg atagtcgcac atcctatgtt atgacttttt 3780
ttactgttat gcttaataaa cgctgcccag atagtcgcac aaactactca ggcaaccctc ataaccctcc 3840
tatggctcct gaatgctggg ctagctccag actgattttt aagaagcggc ggcaaccctc cgctcgagtc 3900
catttacctt ttcctatatt agagaagatg actgattttt ccttctagt ggcagccatc 4020
tagagggccc gtttaacccc gctgatcagc cctcgacttg ggtgcccact ccactgtcct 4080
tgtgtgtttg cctcccccgg tgccttcctt gaccctggaa ggtgtcatt ctattctggg 4140
ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ggcagcctg 4200
gggtgggggt gggcaggaca gcaaggggga ggattgggaa gacaatagca gggcctctag 4260
ggatgcccgt ggctctatgg cttctgaggg gaaagaacc snccntagct ggggtcttag 4320
gggttatccc cagcgccctt gtagcgccgc attaaagcgg cgggggtgtg tgggtacggc 4380
cagcgtgacc gctacacttg ccagcgccct agcgcccgct ccttccgctt tcttcccttc 4440
ctttctcgcc acgttcgccc gctttccccc tcaagctcta aatcggggca tccctttagg 4500
gttccgattt agtgccttac ggcacctcga ccccaaaaaa cttgattagg gtatgggtt 4560
acgtagtggg ccatcgccct gatagacggg ttttcgccc ttgacgttgg agtccaggt 4620
ctttaatagt ggactcttgt tccaaactgg aacccctatc cgggtctatt 4680
ttttgattta taagggattt tggggatttc ggcctatttg ttaaaaaatg agctgattta 4740
acaaaaattt aacgcgaatt aattctgtgg aatgtgtgtc agttagggtg tggaaagtcc 4800
ccaggctccc caggcaggca gaagtatgca aagcatgcat ctcaattagt cagcaaccag 4860
gtgtggaagc tcccagggt cccagcagg cagaagtatg caaagcatgc atctcaatta 4920
gtcagcaacc atagtcccgc cctaactcc gcccatcccg cccctaactc cgccaggtc 4980
cgccattct ccgcccctg gctgactaat tttttttatt tatgcagagg ccgagggcgc 5040
ctctgcctct gagctattcc agaagttagt aggaggttt tttggaggcc taggctttg 5100
caaaaagctc ccgggagctt gtatatccat tttcggatct gatcagcacg tgttgacaat 5160
taatcatcgg catagtatat cggcatagta taatacgaca aggtgaggaa ctaaaaccatg 5220
gccaagttag ccagtgcctt tccggtgctc accgcgcggc acgtcgcggc agcggctcag 5280
ttctggaccg accggtcctg gttctcccgg gacttcgtgg aggacgactt cgccggtgtg 5340
gtccgggacg accgtaccct gttcatcagc gcggtccagg accaggtggg gccgggacaac 5400
acctggcctt ggggtgtggg gcgcgccctg gacgagctgt acgcccagtg gtcggaggtc 5460
gtgtccacga acttccggga cgcctccggg ccggccatga ccgagatcgg cgtggccgag 5520
tggggggcgg agttcgccct gcgcgaccg gccggcaact cggccttcta tgaagggtt 5580
gagcaggact gacacgtgct acgagatttc gattccaccg ccgcttcta tgaagggtt 5640
ggcttcggaa tegttttcgg ggacgcggc tggatgatcc tccagcgccg ggaatcagc 5700
ctggagttct tgcgccaccc caacttggtt attgcagctt tgcatcttag ttgtggttt 5760
aatagcatca caaatttcac aaataaagca tttttttcac tgcattctag ctagagctt 5820
tccaaactca tcaatgtatc ttatcatgtc tgtataccgt cgacctctag aattccacac 5880
gcgtaatcat ggtcatagct gtttcctgtg tgaattgtt atccgctcac

aacatacagag ccggaagcat aaagtgtaaa gcctgggggtg cctaatagagt gagctaactc 5940
acattaattg cggttgcgctc actgcccgcgt ttccagtcgg gaaacctgtc gtgccagctg 6000
cattaatgaa tcggccaacg cgcggggaga ggcgggttgc gtattgggag ctcttccgct 6060
tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac 6120
tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga 6180
gcaaaaaggcc agcaaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 6240
aggctccgcc cccctgacga gcatcacaaa aatcgacgt caagtcagag gtggcgaaac 6300
ccgacaggac tataaagata ccaggcggtt cccctggaa gctccctcgt gcgctctcct 6360
gttccgacce tgccgcttac cggatacctg tccgccttcc tcccttcggg aagcgtggcg 6420
ctttctcaat gctcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg 6480
ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 6540
cttgagtcga acccggttaag acacgactta tcgccactgg cagcagccac tggtaacagg 6600
attagcagag cgaggtatgt aggcgggtgct acagagttct tgaagtggtg gcctaactac 6660
ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga 6720
aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tgggtttttt 6780
gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt 6840
tctacggggg gtgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 6900
ttatcaaaaa ggatcttcac ctgatacctt ttaaattaaa aatgaagttt taaatcaatc 6960
taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 7020
atctcagcga tctgtctatt tccgtctatcc atagtgcct gactccccgt cgtgtagata 7080
actacgatac gggagggcct accatctggc cccagtgctg caatgatacc gcgagaccca 7140
cgctcaccgg ctccagattt atcagcaata aaccagccag cgggaaggcg cgagcgaga 7200
agtggctcct caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga 7260
gtaagtagtt cgcaggttaa tagtttgctg aacgttggtg ccattgctac aggcacatcg 7320
gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga 7380
gttacatgat ccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt 7440
gtcagaagta agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct 7500
cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtaact aaccaagtca 7560
ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc cggcgctcaat acgggataat 7620
accgcgccac atagcagaac tttaaaagtg ctcattcattg gaaaacgttc tcggggcgga 7680
aaactctcaa ggaacttacc gctgttgaga tccagttcga tgtaaccacac tgcgtgacccc 7740
aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg 7800
caaaatgccg caaaaaaggg aataaggcg acacggaaat gttgaatact catactcttc 7860
ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt 7920
gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcga 7980
cctgacgtc 7989

<210> 34

<211> 7607

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: plasmid GRE5-E1-SV40-Hygro

<400> 34

tctagaagat ccgctgtaca ggatgttcta gctactttat tagatccgct gtacaggatg 60
ttctagctac tttattagat ccgctgtaca ggatgttcta gctactttat tagatccgct 120
gtacaggatg ttctagctac tttattagat ccgctgtaca gatgttctag ctactttatt 180
agatcgatct cctggccggt cggggtcaaa aaccaggttt ggctataaaa ggggggtgggg 240
gcgctgtcgt cctcactctc ttccgcatcg ctgtctgcga gggccaggat cgatcctgag 300
aacttcaggg tgagtttggg gacccttgat tggtctttct ttttcgctat tgtaaaattc 360
atgttatatg gagggggcaa agttttcagg gtgttgttta gaatgggaag atgtcccttg 420
tatcaacctg gaccctcatg ataattttgt ttctttcact ttctactctg ttgacaacca 480
ttgtctcctc ttattttctt ttcatcttct gtaacttttt cgttaaactt tagcttgcat 540
ttgtaacgaa tttttaaatt cacttttggt tattgtcag attgtaagta ctttctctaa 600
tcactttttt ttcaaggcaa tcagggtata ttatattgta cttcagcaca gtttttagaga 660
acaaattgta taattaaatg ataaggtaga atatttctgc atataaattc tggctggcgt 720
ggaaatattc ttattggtag aaacaactac atcctgggtc tcatcctgcc ttctcttcta 780
tggttacaaat gatatacact gtttgagatg aggataaaaat actctgagtc caaacggggc 840
ccctctgcta accatgttca tgccctcttc tttttcctac agctcctggg caacgtgctg 900
gttattgtgc tgtctcatca ttttggcaaa gaattagatc taagcttctg cagctcgagg 960

actcgggtcga ctgaaaatga gacatattat ctgccacgga ggtgttatta ccgaagaaat 1020
ggccggccagt cttttggacc agctgacga agaggtagtg gctgataatc tccacctcc 1080
tagccattttt gaaccaccta cccttcacga actgtatgat ttagacgtga cggcccccgga 1140
agatcccaac gaggaggcgg ttttcgagat ttttcccgac tctgtaatgt tggcggtgca 1200
ggaaggggatt gacttactca cttttccgcc ggcggccgggt tctccggagc cgcctcacct 1260
ttcccgccag cccgagcagc cggagcagag agccttgggt ccggtttcta tgccaaacct 1320
tgtaccggag gtgatcgatc ttacctgcca cgaggctggc tttccacca gtgacgacga 1380
ggatgaagag ggtgaggagt ttgtgttaga ttatgtggag caccgccggc acgggtgag 1440
gtcttgtcat tatcaccgga ggaatacggg ggaccagat attatgtgtt cgctttgcta 1500
tatgaggacc tgtggcatgt ttgtctacag taagtgaata ttatgggcag tgggtgtag 1560
agtgggtgggt ttgggtgggt aatttttttt ttaattttta cagttttgtg gtttaaagaa 1620
ttttgtattg tgattttttt aaaaggtcct gtgtctgaac ctgagcctga gcccgagcca 1680
gaaccggagc ctgcaagacc taccgcctgt cctaaaaatgg cgcctgctat cctgagacgc 1740
ccgacatcac ctgtgtctag agaatgcaat atagctgtga ctccggtcct 1800
tctaaccacac ctctgagat acaccgggtg gtcccgctgt gccccattaa accagttgcc 1860
gtgagagttg gtgggcgtcg ccaggctgtg gaatgtatcg aggacttgct taacgagcct 1920
gggcaacctt tggacttgag ctgtaaacgc cccaggccat aagggtgtaa cctgtgattg 1980
cgtgtgtgggt taacgccttt ttgtgtgaa tgagtgtatg taagttaaata aaagggtgag 2040
ataatgttta acttgcatgg cgtgttaaata gggcgggggc ttaaagggtg tataatgcgc 2100
cgtgggctaa tcttggttac atctgacctc atggaggctt gggagtgttt ggaagatttt 2160
tctgtgtgac gtaacttgct ggaacagagc tctaacagta cctcttggtt ttggagtttt 2220
ctgtggggct catcccaggc aaagttagtc tgcagaatta aggaggatta caagtgggaa 2280
tttgaagagc ttttgaaatc ctgtgggtgag ctgtttgatt ctttgaatct gggtcaccag 2340
gcgcttttcc aagagaaggt catcaagact ttggattttt gcgaagaaac ccacccggg 2400
gctgtgtgtg tctttttgag ttttataaag gataaatgga gagcggttgt gagacacaag 2460
gggggggtacc tgctggattt tctggccatg catctgtgga gacggttgt gagacagcag 2520
aatcgctgc tactgttgtc ttccgtccgc ccggcgataa taccgacgga ggagcagcag 2580
cagcagcagg aggaagccag gcggcgccgg caggagcaga gcccatggaa cccgagagcc 2640
ggcctggacc ctccgggaatg aatgtgtgac aggtggctga actgtatcca gaactgagac 2700
gcattttgac aattacagag gatgggcagg ggctaaaggg ggtaaaggag gagcgggggg 2760
cttgtgaggc tacagaggag gctaggaatc tagcttttag cttaatgacc agacaccgtc 2820
ctgagtgtat tacttttcaa cagatcaagg cagctgacca cttactggct gcagccaggg gatgattttg 2880
cgcaagagta tcccatagag gcaaaagggtg caactaggcc agattgcaag tacaagatca 2940
aggaggctat tagggtatat aatatcagg aattgttgct acatttctgg gaacggggcc gaggtggaga 3000
gcaaacttgt ggtatagggtg gcttttagat gtagcatgat aaatatgtgg ccgggggtgc 3060
ttggcatgga cgggggtggtt attatgaatg taagggttac tggccccaat ttagcgggta 3120
cggtttttcc ggccaataacc aaccttatcc tacacgggtg aagcttctat gggtttaaca 3180
atacctgtgt ggaagcctgg accgatgtaa ggggttcggg ctgtgccttt tactgtgct 3240
ggaaggggggt ggtgtgtcgc tctgagggtg actccagggt ggcgccacaat gtggcctccg 3300
ggtgtacctt ggggtacctg gtgaaaagcg tggctgtgat taagcataac atgggtatgtg 3360
actgtgggtg ctcatgcta ggcagggcc tctcagatgc tgacctgtc ggaaggcctg gccagtgttt gatcacctgc 3420
tgaagaccat tcacgtagcc agccactctc gcaaggcctg ggtgttccta ccttaccat 3480
tactgacctg ctgttccttg catttgggtg accccgagag catgtccaag gtgaacctga 3540
gcaatttgag tcacactaag atattgcttg ggaagggtgct gaggtacgat gagacccgca 3600
acgggggtgt tgacatgacc atgaagatct aacatattag gaaccagcct gtgatgctgg 3660
ccagggtgag accctgcgag tgtggcggtg cccgatcact tgggtgctggc ctgcacccgc gctgagtttg 3720
atgtgaccga ggaagataca gattgaggtg ctgaaatgtg tgggcgtggc ttaaggggtg 3780
gctctagcga tgaaggtggg ggtcttatgt agttttgtat ctgttttgca gcagcccgcc 3840
gaaagaatat ataaggtggg tttgatggaa gcattgtgag ctcatatttg acaacgcgca 3900
ccgccatgag caccactcg tttgatggaa tgatgggctc cagcattgat ggtcgccccg 3960
tgccccctg aaactctact acctgacct tcagccgctg cagcctcccg ttcatccgcc 4020
ctgcagcctc cgcccgccgt gcaagcagtg cagcttcccg ggaacttaata cctcctcc 4080
ttgctttcct gagccccgtt caattggatt ctgtgaccg ggtctcctcc cctcccaatg 4140
agttgacggc tcttttgga agcaggttt ctgccctgaa ttggatcaag caagtgtctt 4200
agcagctggt ggatctgcgc aaaccagact cgaattgga tccaattcg atcgatctta 4260
cggtttaaaa cataaataaa aagtgcgaag aatgggttaca tagcatcaca 4320
gctgtctcag ctgactgctt tgcagcttat cattctagtt gtggtttgtc caaactcatc 4380
aatttcacaa ataaagcatt tttttcactg 4440


```

aatgtatctt atcatgtctg gtcgactcta gactcttcg cttcctcgt cactgactcg 4740
ctgcgctcgg tcgttcgggt gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg 4800
ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag 4860
gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc atagggtccg ccccccgtac 4920
gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga 4980
taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt 5040
accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc 5100
tgtaggtatc tcagttcggt gtaggtcggt cgctccaagc tgggctgtgt gcacgaaccc 5160
cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta 5220
agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggat 5280
gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca 5340
gtatgttggt tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct 5400
tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt 5460
acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct 5520
cagtggaaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggtatctc 5580
acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatagtaa 5640
acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta 5700
tttcgttcat cctagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc 5760
ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cagctcacc gggtccgat 5820
ttatcagcaa taaaccagcc agccggaagg gccagcgca gaagtgggtc tgcaacttta 5880
tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag ttccgaggt 5940
aatagtttgc gcaacggttg tgccattgct acaggcatcg tgggtgcacg ctgcgtggt 6000
gggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg atccccatg 6060
ttgtgcaaaa aagcggttag ctcttcggt cctccgatcg ttgtcagaag taagttggcc 6120
gcagtggtat cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc 6180
gtaagtgtct tttctgtgac tggtagtac tcaaccaagt cattctgaga atagtgtatg 6240
cggcgaccga gttgctcttg cccggcgta ataccgggata ataccgccc aaggtatctt 6300
actttaaaaa tgctcatcat tggaaaacgt tcttcggggc gaaaactctc ttcagcatct 6360
ccgctgttga gatccagttc tgggtgagca aaacacaggaa ggcaaaatgc cgcaaaaaag 6420
tttactttca ccagcgtttc tgggtgagca aaacacaggaa ggcaaaatgc cgcaaaaaag 6480
ggaataaggg cgacacggaa atgttgaata ctactactct ttgaatgtat tttagaaaaa 6540
agcattttat aggtttattg totcatgagc ggatacatat cactgacgt ctaagaaaac 6600
aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cagggcccc ttcgtctcgc 6660
attattatca tgacattaac ctataaaaaat aggcgtatca ctcccggaga cggtcacagc 6720
gcgtttcggg gatgacgggtg aaaacctctg acacatgcag ctcccggaga cggtcacagc 6780
ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag ggcgcgtcag cgggtgttg 6840
cgggtgtcgg ggctggctta actatgcggc atcagagcag attgtactga ggtgcacca 6900
tatgcggtgt gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggaaattgta 6960
agcgttaata ttttggttaa attcgcgtta aatttttgtt aaatcagctc attttttaac 7020
caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga gatagggttg 7080
agtgttggtc cagtttgtaa caagagtcca ctattaaaga acgtggactc caactgcaa 7140
gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc ctaatcaagt 7200
tttttggggg cgaggtgccg taaagcacta aatcggaacc cttaaaggag ccccgattt 7260
agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa agcgaaagga 7320
gcggcgctta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac cacaccgcc 7380
gcgcttaatg cggcgctaca gggcgctcc cattcgccat tcaggctcgc caactgttg 7440
gaagggcgat cggtcggggc ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct 7500
gcaaggcgat taagtgggt aacgccaggg ttttcccagt cagcaggtt taaaacgacg 7560
gccagtgaat tgtaatacga ctactatag ggcgaattaa ttcggggg 7607

```

<210> 35
 <211> 11600
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: plasmid MMTV-E2a-SV40-Neo

<400> 35
 gaattccgca ttgcagagat attgtattta agtgccctagc tcgatacaat aaacgccatt 60
 tgaccattca ccacattggg gtgcacctcc aagcttgggc agaaatgggt gaactcccga 120

gagtgtccta cacctagggg agaagcagcc aaggggttgt ttcccaccaa ggacgacccg 180
tctgcgcaca aacggatgag cccatcagac aaagacatat tcattctctg ctgcaaaactt 240
ggcatagctc tgctttgcct ggggtctattg ggggaagtgt cggttcgtgc tcgcagggt 300
ctcacccttg actccttttaa ggggtctattg ggggaagtgt cggttcgtgc tcgcagggt 360
aactcgacct tccctcctgag gcaaggacca cagccaactt cctcttaca ggcgcacatcga 420
ttttgtcctt cagaaataga aataagaatg cttgtctaaa attatatttt taccataaag 480
accaatccaa taggttagatt attagttact atgttaagaa atgaatcatt atcttttagt 540
actattttta ctcaaattca gaagttagaa attgggaatag aaaatagaaa gagacgctca 600
acctcaattg aagaacagggt gcaaggacta ttgaccacag gcctagaagt aaaaaaggga 660
aaaaagagtg tttttgtcaa aataggagac aggtgggtggc aaccagggaac ttatagggga 720
ccttacatct acagaccaac agatgcccc caatggctat aaagtgttat atagatccct cccttttcgt 840
ggggtacagt ggggtacagt acctccttgg tgtatgttgt ctcaagaaga aaaagacgac 900
gaaagactcg ccagagctag acctccttgg tgtatgttgt ctcaagaaga aaaagacgac 960
atgaaacaac aggtacatga ttatattttat cttaggaacag gaatgcactt tggggaaaag 1020
attttccata ccaaggagggt gacagtggct ggactaatag aacattatct tgcaaaaact 1080
catggcatga gttattatga atagccttta ttggcccaac cttgcggttc ccagggtcta 1140
agtaagtttt tggttacaaa ctgttcttaa aacgaggatg tgagacaagt ggtttcctga 1140
cttgggtttg tatcaaaagt tctgatctga gctctgagtg ttctattttt ctatgttctt 1200
ttggaattta taaacttgca acagtcttaa cactttcca gaggggtccc cgcagaccc 1380
atctttttgag ctccgctcgt cacttatect tctctatttc gtgtctgttc 1320
gccatccgt ctccgctcgt cacttatect tctctatttc gtgtctgttc 1320
cggcgacctt gtagaccttg tctctatttc aacggactca ccatagggac caagctagcg cttctcgtcg 1560
ggctatcatc acaagagcgg cctcaaagat ttttggcact tegttagcg gggtggcacg gcaggatagg 1680
cgtccaagac cctcaaagat ttttggcact tegttagcg gggtggcacg gcaggatagg 1680
gcgccctgcc gcaaggccag ctgcttgtcc aaagatgtga taggtggcaa gcacctctgg cactggcaaat 1740
ggatctctgc agttttggaa aaagatgtga taggtggcaa gcacctctgg cactggcaaat 1740
acggggtaga agttgaggcg cgggttgggc tgcgatgtgc taggtggcaa gcacctctgg cactggcaaat 1800
ggtacgcgcg gtgagaatag gtggcgctcg taggcaagc taggcaagc taggcaagc taggcaagc 1860
ggcacatcgc tgcgctcttg caacgcgctg tagtgcgcat gcctttcgtc cccccgcgcg 1980
ttcaacagca cgtcgtctcc tgcgttgcct cctgggtcct gctcgataat cacttctctc 2100
acttgttcc tgcgtctctc gcttcaaaaa gttggtaggc cgttggcggc atcgggtggag 2160
gagcgggtcct ggggtgcttc gacggggaag gtggtaggcg cgttggcggc atcgggtggag 2160
tcctcaagcg ggggtgcttc gacggggaag gtggtaggcg cgttggcggc atcgggtggag 2160
gcggtggttg cgaactcaga gggggcggtt aggtgtcct tcttctcgac tgactccatg 2220
atctttttct gcctatagga gaaggaaatg gccagtccgg aagaggagca gcgcgaaacc 2280
acccccgagc cgcgacgcgg tgcggcgcca cgtcccccaa ccatggagga cgtgtcgtcc 2340
ccgtcccggt gcgcgcgcgc tccccggcg cccccaaaaa agcggtatgag gcggcgatc 2400
gagtccgagg acgaggaaga ctcatcacia gacgcgctgg tgcgcgcac acccagcccg 2460
cggccatcga cctcggcggc ggatttggcc attgcgcca agaggaaaaa gaagcgccct 2520
tctcccaagc ccgagcgccc gccatcacca gaggtaatcg ttggacagcga ggaagaaaga 2580
gaagatgttg cgtacaaaat ggtgggtttc agcaaccac cgggtgcta atcaagcatggc 2640
aaaggaggta agcgcacagt ggcggcgctg aatgaagacg acccagtggc gcgtggtatg 2700
cggacgcaag aggaagagga agagcccagc gaagcggaag gtgaaattac ggtgatgaac 2760
ccgctgagtg tgcgatcgt gtctgcgtgg gagaagggca tggaggctgc gcgcgcgtg 2820
atggacaagt accacgtgga taacgatcta aaggcgaaact tcaaactact gcctgacca 2880
gtggaagctc tggcgccgt atgcaagacc tggctgaacg agggacaccc cgggttgca 2940
ctgaccttc cagcaacaa gacctttgtg acgatgatg ggcgattcct gcaggcgatc 3000
ctgcagctcg ttgcagaggt gacctacaag catcacgagc ccacgggctg cgcgttgtg 3060
ctgcaccgct gcgctgagat cgaaggcgag cttaaagtgc tacacggaag cattatgata 3120
aataaggagc acgtgattga aatggatgtg acgagcgaaa acgggcagcg cgcgctgaag 3180
gagcagtcct gcaaggccaa gatcgtgaag aaccgggtgg gacgcgccct gtcgggcca 3240
tccaacaccg acgcaagggt ctgcgtgcac gacgcgccct gtcgggcca 3300
ggcaagtctt ggcgcatgtt cttctctgaa ggcgcaaaag ctcagggtgg ctcagggtgg 3360
atcaaggctt ttatgcagge gctgtatcct aacgcccaga cggggcacgg tcacctttt 3420
atgccactac ggtgcgagtg caactcaaag cctgggcacg cgcccttttt gggaaggcag 3480
ctaccaaaag tgactccgtt cgccctgagc aacgcgagg acctggacgc ggtatgtatc 3540
tccgacaaga gcgtgctggc cagcgtgcac caccggcg cagctgctgc 3600
aacctgtgt atcgcaactc gcgcgcgcag ggcggaggcc ccaactgcga cttcaagata 3660
tcgggcccgt acctgctaaa cgcgttgggt atggtgcgca gcctgtggag tgaacttct 3720
accgagctgc cgcggatggt tgtgcctgag tttaagtga gcaactaaac ccagtatcgc 3780
aacgtgtccc tgccagtggc gcatagcgat gcgcggcaga accctttga tttttaaacg 3840

gcgcagacgg caaggggtggg ggtaaataat caccgcagag tgtacaaata aaagcatttg 3900
cctttattga aagtgtctct agtacattat ttttcatgt ttttcaagtg acaaaaagaa 3960
gtggcgctcc taatctgcgc actgtggctg cggaagtagg gcgagtggcg ctccaggaag 4020
ctgtagagct gttcctgggt gcgacgcagg gtgggctgta cctggggact gttgagcatg 4080
gagttgggta ccccgtaaat aagggtcatg gtggggtgt gatccatggg agtttggggc 4140
cagttggcaa aggcgtggag aaacatgcag cagaatagtc cacaggcggc cgagttgggc 4200
ccctgtacgc tttgggtgga cttttccagc gttatacagc ggtcggggga agaagcaatg 4260
gcgctacggc gcaggagtga ctgcactca aactggtaaa cctgcttgag tcgctgggtca 4320
gaaaagccaa agggctcaaa gaggtagcat gttttttagt gcgggttcca ggcaaaggcc 4380
atccagtgtg cgccccagc ctgcgcagcc gccgtattga ctatggcgca ggcgagcttg 4440
tgtggagaaa caaagcctgg aaagcgcttg tcataggtgc ccaaaaaata tggcccaaaa 4500
ccaagatctt tgacaatggc tttcagttcc tgctcactgg agcccatggc ggcagctgtt 4560
gttgatgttg cttgcttctt tatgttggg cgttgccggc cgagaagggc gtgcgcaggt 4620
acacggtttc gatgacggcg cggtcgggccc ggtgcacacg gaccacgtca aagacttcaa 4680
acaaaacata aagaagggtg gggatctttg tggatccat ttacttctgt ggtgtgacat 4740
ttacctcagg tcgacctcga gggatctttg tgaaggaaacc ttaacttctgt ggtgtgacat 4800
aattggacaa actacctaca gagattttaa gctctaaggt aaatataaaa tttttaagt 4860
tataatgtgt taaactactg attctaattg tttgtgtatt ttagattcca acctatggaa 4920
ctgatgaatg ggagcagtg tggatgctt ctgactctca acactgttt tgctcagaag 4980
aaatgccatc tagtgatgat gaggctactg ctgactctca acattctact cctccaaaaa 5040
agaagagaaa ggtagaagac cccaaggact ttccttcaga attgctaagt tttttgagtc 5100
atgctgtgtt tagtaataga actcttgctt gctttgctat ttacaccaca aaggaaaaag 5160
ctgcactgct atacaagaaa attatggaaa aatattctgt aaactttata agtaggcata 5220
acagttataa tcataacata ctgttttttc ttactccaca caggcataga gtgtctgcta 5280
ttaataacta tgctcaaaaa ttgtgtacct ttagcttttt aatttgtaaa ggggttaata 5340
aggaatattt gatgtatagt gccttgacta gagatcataa tcagccatac cacatttcta 5400
gaggttttac ttgctttaaa aaacctcccc cactcccc tgaacctgaa acataaaatg 5460
aatgcaattg ttgttgttaa cttgtttatt gcagcttata atggttacaa ataaagcaat 5520
agcatcacia atttcacaaa taaagcattt ttttctactg attctagtgt tggtttgtcc 5580
aaactcatca atgtatctta tcatgtctgg atccggctgt ggaatgtgtg tcagttaggg 5640
tgtggaaagt ccccgagtc cccagcaggc agaagtatgc aaagcatgca tctcaattag 5700
tcagcaacca ggtgtggaag gtccccaggg tccccagcag gcagaagtat gcaaagcatg 5760
catctcaatt agtcagcaac cccctaaact cgcccatccc gcccctaact 5820
ccgcccagtt ccgcccattc tccgccccat ggctgactaa ttttttttat ttatgcagag 5880
gcccaggccg cctcgccctc tgagctattc cagaagtagt gaggaggctt ttttggaggc 5940
ctaggctttt gcaaaaagct tcacgctgcc gcaagcactc agggcgcaag ggtgctaaa 6000
ggaagcggaa cacgtagaaa gccagtcgag gaaacgcgtg atgaatgtca 6060
gtactgggc tatctggaca agggaaaaac caagcgcaaa ctgaccccg gtgacttgca 6120
gtgggcttac atggcgatag ctagactggg cggttttatg gagaaagcag gtgacttgca 6180
tgccagctgg ggcgcccctc ggtaagggtg ggaagccctg caaagttaac tggatggctt 6240
tcttgccgcc aaggatctga tggcgaggg gatcaagag tgatcaagag acaggatgag 6300
gatcgtttcg catgattgaa caagatggat tgcacgcagg ttctccggcc gcttgggtgg 6360
agaggctatt cggctatgac tgggcacaa agacaatcgg ctgctctgat gccgccgtgt 6420
tccggctgtc agcgcagggg cgcccgggtc tttttgtcaa gaccgacctg tccggtgccc 6480
tgaatgaact gcaggacgag gcagcgcgcc tatcgtggct ggccacgacg ggcgttccct 6540
gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta ttgggcgaag 6600
tgccggggca ggtctcctg tcatctcacc ttgctcctgc cgagaaagta tccatcatgg 6660
ctgatgcaat gcggcggtg catacgttg atccggctac cgtctctgtc gaccaccaag 6720
cgaaacatcg catcgagcga gcacgtactc ggttggaagc cgtctctgtc gatcaggatg 6780
atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg ctcaaggcgc 6840
gcatgcccga cggcgaggat ctgcgtctga cccatggcga tgctgcttg ccgaatatca 6900
tgggtgaaaaa tggccgcttt totggattca tcgactgtgg ccggtgggt gtggcggaac 6960
gctatcagga catagcgttg gctaccctg atattgtga agagcttggc ggcgaatggg 7020
ctgacgctt cctcgtgctt tacggtatcg ccgctcccga ttcgacgcgc atcgcttct 7080
atcgcttct tgacgagttc ttctgagcgg gactctgggg ttccaccgcc gccttctatg aaaggttggg 7200
gacgcccac ctgccatcac gagatttcga ttccaccgcc cagcgcgggg atctcatgct 7260
cttcggaatc gttttccggg acgcgggctg gatgacctc tggttcagct gctgcctgag 7320
ggagttcttc gcccaccccg ggctcgatcc cctcgcgagt gtcggcatcc aggaaccag 7380
gctggacgac ctgcggagt tctaccggca actgcaggag tggggaggca cgatggccgc 7440
cagcgctat ccgcatcc atgccccga acttctgtgg tgtgacataa ttggacaaa 7500
tttggtccc gatctttgtg tctaaggtaa atataaaatt ttaagtgtg 7560
tacctacaga gatttaaagc

aactactgat tctaattggt tgtgtatgtt agattccaac ctatggaact gatgaatggg 7620
agcagtgggt gaatgccttt aatgaggaaa acctgttttg ctacagaaga atgccatcta 7680
gtgatgatga ggctactgct gactctcaac attctactcc tccaaaaaag aagagaaaag 7740
tagaagaccc caaggacttt ccttcagaat tgctaagttt tttgagtcac gctgtgttta 7800
gtaatagaac tcttgcttgc tttgtctatt acaccacaaa ggaaaaagct gcactgctat 7860
acaagaaaat tatggaaaaa tattctgttaa cctttataag taggcataac agttataatc 7920
ataacatact gttttttctt actccacaca ggcataagat gtctgctatt aataactatg 7980
ctcaaaaatt gtgtaccttt agctttttta tttgtaaaag ggttaataag gaataattga 8040
tgtatagtgc cttgactaga gatcataatc agccatacca catttgtaga ggttttactt 8100
gcttttaaaa acctcccaca cctccccctg aacctgaaac ataaaaatgaa tgcaattggt 8160
gttggttaact tgtttattgc agcttataat ggttacaaat aaagcaatag catcacaaat 8220
ttcacaaaata aagcattttt ttcactgcat tctagtgtg gtttgtccaa actcatcaat 8280
gtatcttatac atgtctggat ccccgaggag ctccctctgtg tccctcataaa ccctaaccctc 8340
ctctacttga gaggacattc caatcatagg ctgcccctcc accctctgtg tccctcgtgt 8400
aattaggtca ctttaacaaa aggaattgg gaagtcctt ccactgctgt gttccagaag gctttctaag 8460
ggtaattttta aaatatctgg gaagtcctt ccactgctgt gttccagaag gctttgcaca 8520
cagccacaaa atgtcaacag cagaacata caagctgtca gctttgcaca agggcccaac 8580
accctgctca tcaagaagca ctgtgttagg ttccaaaata tctagtgtt tcatTTTTac ttggatcagg 8640
atTTTcccc cctgtgttag cctactggat aagcattatc tgggggttaca gtttgagcag gatatttgg 8700
aaccagcac ctgtcaactg tagcattttt cgaaggttcc caaccaacag caaaaaaatg 8760
catctgctga gctaacacac cctgcaactc accattttca aacctcagtt ttaacagtaa 8820
cctgtagtgt gcttaacacac ccttgaatgg cttttccagc accattttca aacctcagtt ttaacagtaa 8880
aaaatttgac ccttgaatgg cttttccagc accattttca aacctcagtt ttaacagtaa 8940
atgcaagttt aacatagcag ttaccccaat aacctcagtt ttaacagtaa 9000
catcaaaaata tttccacagg ttaagtctc cctattttta atttaaatta ggcaaggaa 9060
cgaaagggcc tctgtatag cctattttta atttaaatta ggcaaggaa 9120
tagacgtcag gtggcaactt tccggtcatg agacaataac cctgataaat 9180
taaatacatt ggaagagtat gagtattcaa ctttccctg tgggtgaaagt 9240
tattgaaaaa ggaagagtat gagtattcaa ctttccctg tgggtgaaagt 9300
gcggcatttt ggggtgcaag agtgggttac atcgaaactg atctcaacag cggttaagatc 9360
gaagatcagt tgggtgcaag agtgggttac atcgaaactg atctcaacag cggttaagatc 9420
cttgagagtt ttcgccccga agaactgttt ccaatgatga gcacttttaa agttctgcta 9480
tgtggcgagg tattatcccg tgttgacgac ccaatgatga gcacttttaa agttctgcta 9540
tattctcaga atgacttggg tgagtactca ccagtcacag aactcggtcg ccgcatacac 9600
atgacagtaa gagaattatg cagtgtgccc ataaccatga gtgataacac tacgggatggc 9660
ttactcttga caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg 9720
gatcatgtaa ctgccttga tgcagcaatg gaggagctga atgaagccat accaaacgac 9780
gagcgtgaca ccacgatgcc cgggcaacaa ttaatagact ggatggaggc ggataaagtt 9840
gaactactta ctctagcttc cggccttccg gctggctggt ttattgctga taagtctgga 9900
gcaggaccac ttctgcgctc gggccttccg gctggctggt ttattgctga taagtctgga 9960
gccggtgagc gtgggtctcg cgggtatcatt gctggctggt ttattgctga taagtctgga 10020
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 10080
atcgctgaga taggtgcctc actgattaaag cattggtaac tgtcagacca agtttactca 10140
tatatacttt agattgattt aaaaacttcat ttttaattta aaaggatcta ggtgaagatc 10200
ctttttgata atctcatgac caaaatccct taacgtgagt ttctgttcca ctgagcgtca 10260
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaactcgc 10320
tgcttgcaaa caaaaaaacc accgctacca gcggtgggtt gtttgccgga tcaagagcta 10380
ccaactcttt ttccgaaggt aactggcttc aagagagcgc agataccaaa tactgtcctt 10440
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc 10500
gctctgctaa tctgtttacc agtggctgct gccagtggcg ataatgctg tcttaccggg 10560
ttggactcaa gacgatagtt accggataag ggcagcggg cgggctgaac ggggggttcg 10620
tgacacagc ccagcttggg gcgaacgacc tacaccgaac tgagatacct acagcgtgag 10680
ctatgagaaa gcgcccagct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc 10740
aggggtcgaa caggagagcg caccgagggg cttccagggg gaaacgcctg gtatctttat 10800
agtctgtcgc ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctctgaggg 10860
gggcgagacc tatggaaaaa cgccagcaac ggggcctttt tacggttctt ggccttttgc 10920
tggccttttg ctcacatgtt ctttccctgc ttatcccttg attctgtgga taaccgtatt 10980
accgcctttg agtgagctga taccgctcgc cgcagccgaa cagcgagtc cagcgagtc 11040
gtgagcgagg aagcggaaga gcgctgatg cggatatttt tccttacgca tctgtgcgg 11100
atttcacacc gcatatggtg cactctcagt acaatctgct ctgatgccgc atagttaagc 11160
cagtatctgc tccctgcttg tgtgttggag gtcgctgagt agtgcgcgag 11220
gctacaacaa ggcaaggctt gaccgacaat tgcatgaaga atctgcttag ggttaggcgt 11280

tttgcgctgc ttccgcatgt acggggccaga tatacgcgta tctgagggga ctagggtgtg 11340
tttagcgcaa aagcggggct tcggttgtag gcggttagga gtccctcag gatatagtag 11400
tttcgctttt gcatagggag ggggaaatgt agtcttatgc aatacacttg tagtcttgca 11460
acatggtaac gatgagttag caacatgcct tacaaggaga gaaaaagcac cgtgcatgcc 11520
gattggtgga agtaagggtg tacgatcgtg ccttattagg aaggcaacag acgggtctga 11580
catggattgg acgaaccact 11600

<210> 36
<211> 53
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 36
gtcactcgag gactcggctg actgaaaatg agacatatta tctgccacgg acc 53

<210> 37
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 37
cgagatcgat cacctccggt acaaggtttg gcatag 36

<210> 38
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 38
catgaagatc tggaagggtg tgaggtacga tgagacc 37

<210> 39
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 39
gcgacttaag cagtcagctg agacagcaag acacttgctt gatccaaatc c 51

<210> 40
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 40
cacgaattcg tcagcgcttc tcgtcgcgtc caagaccc 38

<210> 41
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 41
caccgccgggg aggcggcgccg gacgggggacg gg

32

<210> 42
<211> 7231
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: plasmid pDV80

<400> 42
ctgctccctg cttgtgtgtt ggaggtcgct gtagtagtgcg cgagcaaaat ttaagctaca 60
acaaggcaag gcttgaccga caattgcatg aagaatctgc ttaggggttag gcgttttgcg 120
ctgcttcgcg atgtacgggc cagatatacg cgttgacatt gattattgac tagttattaa 180
tagtaataca ttacgggggtc attagttcat agcccatata tggagttccg cgttacataa 240
cttacggtaaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 300
atgacgtatg ttcccatagt aacgcccaata gggactttcc attgacgtca atgggtggac 360
tatttacggt aaactgcccc cttggcagta catcaagtgt atcatatgcc aagtacgccc 420
cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta 480
tgggactttc ctacttggca gtacatctac gtatttagtca tcgctattac catgggtgatg 540
cggtttttggc agtacatcaa tgggcgtgga tagcgggttg actcacgggg atttccaagt 600
ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca 660
aaatgtcgta acaactccgc cccattgacg caaatggcg gtaggcgtgt acggtgggag 720
gtctatataa gcagagctct ctggctaact agagaaccca ctgcttactg gcttatcgaa 780
attaatacga ctcaactatg ggagacccaa gctggctagc gtttaaactt aagcttggt 840
ccgagctcgg atccactctc ttccgcatcg ctgtctcgca gggccagctg ttgggggtgag 900
tactccctct gaaaagcggg catgacttct gcgctaagat tgctagtttc caaaaaacgag 960
gaggatttga tattcaactg gcccgcggtg atgcctttga ggggtggccg atccatctgg 1020
tcagaaaaga caatcttttt gttgtcaagc ttggtggcaa acgaccgta gagggcggtg 1080
gacagcaact tggcgatgga gcgcagggtt tgggttttgt cgcgatcggc gcgtccctg 1140
gccgcgatgt ttagctgcac gtattcgcgc gcaacgcacc gccattcggg aaagacgggtg 1200
gtgcgctcgt cgggcaccag gtgcacgcgc caaccgcgtt tgtgcagggt gacaagggtca 1260
acgctgggtg ctacctctcc gcgtaggcgc tcgttgggtc agcagaggcg gccgcccttg 1320
cgcgagcaga atggcggtag ggggtctagc tgcgtctcgt cccgggggggtc tgcgtccacg 1380
gtaaaagacc cgggcagcag gcgcgcgtcg aagtagtcta tcttgcaccc ttgcaagtct 1440
agcgctgct gccatgcgcg ggcggcaagc gcgcgctcgt atgggttgag tgggggaccc 1500
catggcatgg ggtgggtgag cgcggaggcg tacatgccgc aaatgtcgta aacgtagagg 1560
ggctctctga gtattccaag atatgtagg tagcatcttc caccgcggat gctggcgcg 1620
acgtaatcgt atagttcgtg cgagggagcg aggaggtcgg gaccgaggtt gctacggggc 1680
ggctgctctg ctcggaagac tatctgcctg aagatggcat gtgagttgga tgatatggtt 1740
ggacgctgga agacgttgaa gctggcgtct gtgagacct cccggtcacg cacgaaggag 1800
gcgtaggagt cgcgcagctt gttgaccagc tcggcggtga cctgcacgtc tagggcgcg 1860
tagtccaggg tttccttgat gatgtcatac ttatcctgtc cctttttttt ccacagctcg 1920
cggttgagga caaactcttc gcgggtctttc cagtactctt ggatcggaat cccgtcgggc 1980
tccgaacgag atccgtactc gcgcgcgag ggacctgagc gactccgcat cgaccggatc 2040
ggaaaacctc tcgagaaagg cgtctaacca gtcacagtcg caagatccaa gatgaagcgc 2100
gcccgcccca gcgaagatga cttcaacccc gtctaccct atggctacgc gcggaatcag 2160
aatatccctc tccctactcc cccctttgtc tcctccgatg gattcaaaaa cttccccctc 2220
ggggtactgt cactcaaact ggctgatcca atcaccatta ccaatgggga tgtatccctc 2280
aagggtgggag gtgggtctcac tttgcaagat ggaagcctaa ctgtaaaccc taagggtcca 2340
ctgcaagtta atactgataa aaaacttgag cttgcatatg ataatccatt tgaaagtagt 2400
gctaataaac ttagtttaaa agtaggacat ggattaaaaa tattagatga aaaaagtgtc 2460

gcgggggttaa aagatttaaat tggcaaaactt gtgggttttaa caggaaaagg aataggcact 2520
gaaaatttag aaaaatacaga tggtagcagc agaggaattg gtataaatgt aagagcaaga 2580
gaagggttga catttgacaa tgatggatac ttggtagcat ggaacccaaa gtatgacacg 2640
cgcacacttt ggacaacacc agacacatct ccaaactgca caattgctca agataaggac 2700
tctaaactca ctttggtact tacaaagtgt ggaagtcaaa tattagctaa tgtgtctttg 2760
atttgggtcg caggaaagta ccacatcata aataataaga ggaagtgttt tagacaactc aaatcttggg 2820
tttactatta aactgctatt taataagaac tccaatgttt cgacagctta tgaaaaagca 2880
aaagcttatt ggaacttttag aagtggaaat ccaaaaccca gtaattctaa aaaatatgca 2940
atttggtttta tgcctaattt ggtagcgtat gggtggaaaac ctgatcagcc agcagtcatt 3000
agagacatag tttatggaaac tatatatctt gaatactcta tcacatttaa ctttagttgg 3060
aaaactacct ttaaccaaga aactggatgt accactctt ttaccttctc ctatatgtgc 3120
tccaaaacct atgaaaatgt tgaatttgaa accagctctt ttaaacccgc tgatcagcct 3180
caagaatgaa agagcgggccc ctcgagtcta gaggggccgt ttaccccggt ccttcttga 3240
cgactgtgccc ttctagtggc cagccatctg ttgtttgccc ctcccccggt gcacgcgatt 3300
ccctgggaagg tgcactccc actgtccttt cctaataaaa tgaggaaatt gcacgcgatt 3360
gtctgagtag gtgtcattct attctggggg gtgggggtggg gcaggacagc aagggggagg 3420
atttgggaaga caatagcagg catgtcgggg atgcggtggg ctctatggct tctgaggcgg 3480
aaagaaccag ctggggctct aggggggtat ccacgcgcc ctgtagcggc gcattaaagg 3540
cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctacgcggcc 3600
ctcctttcgc tttcttccct tcctttctcg ccacgttctg cggctttccc cgtcaagctc 3660
taaatcgggg catcccttta gggttccgat ttagtgtctt acggcacctc gaccccaaaa 3720
aacttgatta gggtagtgt tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc 3780
ctttgacggt ggagtccacg ttctttaata tgggactctt gttccaaact ggaaccaac 3840
tcaaccctat ctcggtctat tcttttgatt ataagggat tttggggatt tggcctatt 3900
ggttaaaaaa tgaactgatt taacaaaaat ttaacgcgaa ttaattctgt ggaatgtgtg 3960
tcagttaggg tgtggaaagt cccagggctc ccagggcagg cagaagtatg caaagcatgc 4020
atctcaatta gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta 4080
tgcaaaagcat gcatctcaat tagtcagcaa ccatagtccc gccctaact ccgcccattc 4140
cgcccctaac tccgcccagt tccgcccatt ctccgcccga tgggtgacta atttttttta 4200
tttatgcaga ggccgaggcc gctctgcct ctgagctatt ccagaagtag tgaggaggct 4260
tttttgagg cctaggcttt tgcaaaaagc tcccgggagg ttgtatatcc attttcggat 4320
ctgatcagca cgtgttgaca attaatcctc gccatagat atcggcatag tataatacga 4380
caaggtgagg aactaaacca tggccaagtt gaccagtgcg gttccggtgc tcaccgcggc 4440
cgacgtcgcc ggagcgggtc agttctggac cgaccggctc ggggttctcc gggacttcgt 4500
ggaggacgac ttccgcccgtg tgggtccggga cgacgtgacc ctggtcatca gcgcgggtcca 4560
ggaccaggtg gtgcccggaca acaccctggc ctgggtgtgg gtgcccggcc tggacgagct 4620
gtaccgagag tgggtcggagg tcgtgtccac gaacttccgg gacgcctccg ggccggccat 4680
gaccgagatc ggcgagcagc cgtggggggc ggtacacgtg ctacgagatt tgcattccac 4740
ctgcgtgcac ttctgtggccg agggagcagga ctgacacgtg aatcgttttc cgggacgccc gctggatgat 4800
cgccgccttc tatgaaagggt tgggcttcgg aatcgttttc cccaacttgt ttattgcagc 4860
cctccagcgc ggggatctca tgctggagtt cttcggccac ccaaaatttg catttttttc 4920
ttataatggt tacaaaataa gcaatagcat cacaattttc aaaaataaag tctgtatacc 4980
actgcattct agttgtgggt tgtccaaact catcaatgta tcttatcatg tctgtatacc 5040
gtcgacctct agctagagct tggcgtaatc atgggtcatag ctgttttctg tgtgaaattg 5100
ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgtg aagcctgggg 5160
tgccaatga gtgagctaac tcacattaat tgcgttgccg tcaactgccc ctttccagtc 5220
gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgccgaggga gaggcgggtt 5280
gcgtattggg cgctcttcg cttcctcgtc cactgactcg ctgcgctcgg tcgttcggct 5340
gcggcgagcg gtatcagctc actcaaaggc ggttaatacgg ttatccacag aatcaggggg 5400
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 5460
cgcggtgctg gcgtttttcc ataggctccg ccccccgtgac gagcatcaca aaaatcgacg 5520
ctcaagtccg aggtggcgaa acccgacagg actataaaga accggatacc taccaggcgt 5580
aagctccctc gtgcgctctc cgtttccgac cctgcccgtt accggatacc tgtccgctt 5640
tctcccttcg ggaagcgtg cgctttctca atgctcacgc tgtaggtatc tcagttcgg 5700
gtaggtcgtt cgctccaagc tgggctgtgt atgctcacgc cccgttcagc cccgacgact 5760
cgccctatcc ggtaactatc gtcttgagtc caaccgggta agacacgact tatcgccact 5820
ggcagcagcc actggtaaca ggattagcag agcgaggat gtaggcgggt ctacagagtt 5880
cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggtg tctgcgctct 5940
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgcacggca aacaaaccac 6000
cgctggtagc ggtgggtttt ttgtttgcaa acgcgcagaa aaaaaggatc 6060
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg 6120
ttaagggatt ttggctcatg gattatcaaa aaggatcttc acctagatcc ttttaaat 6180

```

aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 6240
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 6300
ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc 6360
tgcaatgata cgcgagacc cacgtcacc ggctccagat ttatcagcaa taaaccagcc 6420
agccggaagg gccgagcgca gaagtgggtcc tgcaacttta tccgcctcca tccagtctat 6480
taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt 6540
tgccattgct acaggcatcg tgggtgtcacg ctcgctggtt ggtatggctt cattcagctc 6600
cgggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcgggttag 6660
ctccttcggt cctccgatcg ttgtcagaag taagtggcc gcagtgttat cactcatggg 6720
tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac 6780
tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg 6840
cccggcggtca atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat 6900
tggaanaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc 6960
gatgtaaccc actcgtgcac ccaactgac ttcagcatct tttactttca ccagcgtttc 7020
tggttgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa 7080
atgttgaata ctcatactct tcctttttca atattattga agcatttatc agggttattg 7140
tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttcccg 7200
cacatttccc cgaaaagtgc cacctgacgt c 7231

```

<210> 43

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 43

tgtcttggat ccaagatgaa gcgcgcccgc cccagcgaag atgacttc

48

<210> 44

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 44

aaacacggcg gccgctcttt cattcttg

28

<210> 45

<211> 17

<212> PRT

<213> Ad 37 N-terminus

<400> 45

```

Met Ser Lys Arg Leu Arb Val Glu Asp Asp Phe Asn Pro Val Tyr Pro Tyr
  1           5           10           15

```

<210> 46

<211> 6

<212> PRT

<213> artificial sequence

<400> 46

```

Lys Arg Ala Arg Pro Ser
  1           5

```

<210> 47

<211> 17

<212> PRT
<213> Ad5 modified N-terminus

<400> 47
Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro Tyr
1 5 10 15

<210> 48
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 48
ggatccatgg gatacttggt agca 24

<210> 49
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 49
gcaactcgag tcattcttgg gcaatatagg 30

<210> 50
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 50
cgcgctgact cttaaggact agtttc 26

<210> 51
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 51
gcgcttaatt aacatcatca ataataacc ttatattt 37